MESH INDEPENDENT CONVERGENCE OF MODIFIED INEXACT NEWTON METHODS FOR SECOND ORDER NONLINEAR PROBLEMS

A Dissertation

by

TAEJONG KIM

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2006

Major Subject: Mathematics

MESH INDEPENDENT CONVERGENCE OF MODIFIED INEXACT NEWTON METHODS FOR SECOND ORDER NONLINEAR PROBLEMS

A Dissertation

by

TAEJONG KIM

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Joseph E. Pasciak
Committee Members,	Bojan Popov
	Wolfgang Bangerth
	Vivek Sarin
Head of Department,	Albert Boggess

May 2006

Major Subject: Mathematics

ABSTRACT

Mesh Independent Convergence of Modified Inexact Newton Methods for Second Order Nonlinear Problems. (May 2006)

Taejong Kim, B.S., Yonsei University, Korea;

M.S., Yonsei University, Korea

Chair of Advisory Committee: Dr. Joseph E. Pasciak

In this dissertation, we consider modified inexact Newton methods applied to second order nonlinear problems. In the implementation of Newton's method applied to problems with a large number of degrees of freedom, it is often necessary to solve the linear Jacobian system iteratively. Although a general theory for the convergence of modified inexact Newton's methods has been developed, its application to nonlinear problems from nonlinear PDE's is far from complete. The case where the nonlinear operator is a zeroth order perturbation of a fixed linear operator was considered in the paper written by Brown *et al.*.

The goal of this dissertation is to show that one can develop modified inexact Newton's methods which converge at a rate independent of the number of unknowns for problems with higher order nonlinearities. To do this, we are required to first, set up the problem on a scale of Hilbert spaces, and second, to devise a special iterative technique which converges in a higher order Sobolev norm, i.e., $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ with $0 < \alpha < 1/2$. We show that the linear system solved in Newton's method can be replaced with one iterative step provided that the initial iterate is close enough. The closeness criteria can be taken independent of the mesh size.

In addition, we have the same convergence rates of the method in the norm of $H_0^1(\Omega)$ using the discrete Sobolev inequalities.

To God and my family

ACKNOWLEDGMENTS

I first thank God for giving me a chance to study and for helping to finish my Ph.D. degree Texas A&M University.

I give great thanks to my adviser, Dr. Joseph E. Pasciak, for his leading, help, encouragement, and giving insights to do research throughout my graduate studies. I thank him for giving interesting lessons and for introducing new knowledge. Without his academic knowledge and editorial advice, this dissertation could not have been completed.

I thank all members of my advisory committee for their numerous comments on my research and encouragement. I thank Dr. Bojan Popov for his kindness and for his willingness to help. I thank Dr. Wolfgang Bangerth for giving his interest to my research topics. I thank Dr. Vivek Sarin for serving as a member of my advisory committee and for giving interesting comments on my research.

I thank Dr. James Bramble for sharing his inspiring ideas and for teaching valuable courses, such as multigrid methods and domain decomposition methods. I thank Dr. Raytcho Lazarov for his kindness, encouragement, love, and help in my graduate studies.

I thank Dr. Jean-Luc Guermond, Dr. Jay Walton, Dr. Thomas Schlumprecht, Dr. Paulo Lima-Filho, Ms. Stewart Monique, and all other staff in the Department of Mathematics. I also thank my friends in mathematic department, Dukjin Nam, Seungil Kim, Dylan Copeland, Paul Dostert, Veselin Dobrev and Dimitar Trenev for sharing and discussing many useful ideas.

I give deepest thanks to my family, especially to my wife Junghwa Do, for their endless love, encouragement, and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The purpose of this dissertation is to provide convergence estimates for modified inexact Newton methods applied to second order nonlinear problems where the nonlinearity appears in the coefficient of the highest order derivatives. Specifically, we consider the model problem:

$$\begin{aligned} -\operatorname{div}(k(u,x)\boldsymbol{\nabla} u) + \boldsymbol{c}(u,x)\cdot\boldsymbol{\nabla} u + b(u,x)u = f, \quad x \in \Omega, \\ u(x) = 0, \quad x \in \partial\Omega. \end{aligned} \tag{NP}$$

Here k, b, and $\boldsymbol{c}$ are smooth functions of u and x on $\bar{\Omega}$ which is a bounded polygonal domain in $\mathbb{R}^d$, $d = 2$ or 3. In addition, k is bounded away from zero. In this dissertation, we shall not consider discontinuous coefficients. For brevity, in the analysis to follow we will assume $b = 0$ and $\boldsymbol{c}$ to be independent of u. It will be clear that everything we do carries over to the more general form of the problem given in (NP).

Newton's method is very attractive since its convergence is fast if the initial guess is sufficiently close to the solution. However, often to obtain the solution to the linear Jacobian system becomes too expensive when the number of unknowns is large. Inexact Newton or Newton Krylov methods instead replace the Jacobian solve by a fixed number of steps in a preconditioned iterative procedure.

For several decades, inexact Newton methods have been commonly used for solving nonlinear problems. Dembo *et al.* [22] introduced the inexact Newton method and obtained local convergence results. Later, the affine invariant conditions for inexact Newton methods have been considered by several authors (see, e.g., [1, 24, 35, 42]) to

---

The dissertation model is *SIAM Journal on Numerical Analysis.*

deal with problems whose Jacobian matrices are ill–conditioned. Also, in [4, 23, 24, 25], the authors presented the applications of globalization techniques of the Newton method to inexact Newton methods. Recently, several authors have applied inexact Newton methods in different contexts (see, e.g., [14, 15, 37]).

Estimates which give rise to a uniform rate of iterative convergence for inexact Newton's method applied to (NP) in the case when the nonlinearities are restricted to the zeroth order term were provided in [14] (specifically, $k(u,x) \equiv k(x)$ and $\boldsymbol{c} \equiv 0$). This restriction enabled them to use the $L^2(\Omega)$-norm for the residuals. Iterative methods with residual convergence in $L^2(\Omega)$ can be constructed in the case of full elliptic regularity by multilevel methods as discussed in [14]. Unfortunately, when the coefficients of the higher derivatives involve the discrete solution, full elliptic regularity no longer holds.

We provide a general theorem for the analysis of inexact Newton methods which is a variant of those given in, e.g., [14, 22, 24]. The application of this theorem requires an iterative scheme which reduces the error in a norm which is related to the stability properties of the partial differential equation. The natural norms which have been used in the stability analysis of nonlinear partial differential equations (PDE's) typically involve the Sobolev space $W_0^{1,p}(\Omega)$ for $p > d$ and its dual (see, e.g., [13, 17]). When this is put into our inexact Newton framework, one requires an iterative scheme which reduces the error in $W_0^{1,p}(\Omega)$. Unfortunately, to the best of our knowledge, efficient fixed step iterative techniques which guarantee a reduction in this norm are not known. All of the popular techniques, e.g., multigrid and domain decomposition, for analyzing the iterative convergence for the discrete systems resulting from approximations to PDE's are based on the use Hilbert space and give rise to reductions in the corresponding energy norm.

To deal with this problem, in Chapter IV we analyze the PDE and inexact

Newton's method in the scale of Sobolev norms $H^{1+\alpha}(\Omega)$, for $0 < \alpha < 1/2$ (see Remark B.1 in Chapter II). For two-dimensional problems, this norm coerces the norm in $L^\infty(\Omega)$ by the Sobolev inequality (see Theorem A.4 in Chapter II). To apply our abstract convergence results for the inexact Newton method, we are required to analyze both the continuous and discrete nonlinear problems in these norms. We then get that a uniform convergence rate for the modified inexact Newton method will be achieved provided that one uses an iterative procedure which is a reduction operator in a discrete norm equivalent to the norm in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$. We shall develop such an iterative method based on the work in [7]. In Appendix A, we shall see that the W–cycle multigrid method is also an iterative scheme which reduces the error in a discrete norm equivalent to the norm in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$.

In three-dimensional cases, we first analyze the continuous nonlinear problems in the Sobolev space $W_0^{1,p}(\Omega)$, not in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ because the norm in the latter space does not coerce the norm in $L^\infty(\Omega)$. Next we obtain mesh independent convergence rates for the inexact Newton method in the norm of $H_0^1(\Omega)$ using discrete Sobolev inequalities (see Theorem B.4). Unfortunately, the initial iterate of the method must be close to the exact solution, and the closeness depends on the mesh size.

The dissertation is organized as follows. In Chapter II, we first introduce the Sobolev spaces and several useful inequalities in these spaces. Then the finite element spaces and properties of projections, such as nodal interpolation, the L^2 projection, and the elliptic projection, are presented. Several technical lemmas and the preconditioned conjugate gradient method are also discussed in this chapter. In Chapter III, we describe the modified inexact Newton method and its convergence theorem. In Chapter IV, we introduce a nonlinear problem and analyze the problem in the Hilbert space $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$. The existence of a discrete solution to the problem and an iterative method which reduces the error in a discrete norm equivalent to the norm in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ are also presented in this chapter. In addition, mesh independent convergence rates and some numerical results for the modified inexact Newton method with the PCG method are given there. In Chapter V, we analyze the nonlinear problem in the Hilbert space $H_0^1(\Omega)$ when Ω is the subset of $\mathbb{R}^d$, $d = 2$ or 3. The uniform convergence rates of the modified inexact Newton method are discussed. Finally, there are conclusions and future works in Chapter VI.

CHAPTER II

PRELIMINARIES

In these preliminaries, we introduce the basics for Sobolev spaces, finite element spaces, a few technical lemmas, and the preconditioned conjugate gradient (PCG) method. First, we present some definitions, notations and embedding theorems for Sobolev spaces. Then we describe finite element spaces and introduce operators, such as the Lagrange interpolation operator, the L^2 projection and elliptic projections, which map from a Sobolev space to a finite element space. Next, several technical lemmas which play important roles in this disseration are introduced. In the last section, the PCG method and its convergence rate are presented. We will apply the PCG method to find the correction in the inexact Newton method.

A. Sobolev spaces

There are many references which develop Sobolev spaces and their properties (see, e.g., [13, 19, 27, 28, 43]). For completeness, we shall define Sobolev spaces. Sobolev embedding theorems and various Sobolev inequalities will be given in this section.

1. Definitions and notations

Let Ω be a bounded and open subset in $\mathbb{R}^d$, for $d = 2,\, 3$, whose boundary $\partial\Omega$ is Lipschitz continuous (see, e.g., Definition 1.44 in [13]). To define partial derivatives, we introduce the multi–index notation. Let α be a multi–index whose components α_i, $i = 1, \ldots, d$, are non–negative integers. The length of α is defined by

$$|\alpha| = \sum_{i=1}^{d} \alpha_i.$$

For $\phi \in C^\infty(\Omega)$, denote

$$D^\alpha \phi = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_d}\right)^{\alpha_d} \phi.$$

Here $x = (x_1, \cdots, x_d)$ is the coordinate system in $\mathbb{R}^d$. Note that $|\alpha|$ is called the order of this derivative.

Now, we shall define the Sobolev spaces. The corresponding norms involve the weak derivatives of functions.

Definition A.1 (Sobolev spaces). For a given non–negative integer k and $1 \le p \le \infty$, we define the Sobolev space, $W^{k,p}(\Omega)$, as the set of all functions $f \in L^p(\Omega)$ whose weak derivatives, $D^\alpha f$, are $L^p(\Omega)$ for all $|\alpha| \le k$. The corresponding Sobolev norm is defined by

$$\|f\|_{k,p} = \begin{cases} \left(\sum_{|\alpha| \le k} \|D^\alpha f\|^p_{L^p(\Omega)} \right)^{1/p}, & \text{if } 1 \le p < \infty, \\ \max_{|\alpha| \le k} \|D^\alpha f\|_{L^\infty(\Omega)}, & \text{if } p = \infty. \end{cases}$$

Furthermore, we define the Sobolev semi–norm

$$|f|_{k,p} = \begin{cases} \left(\sum_{|\alpha| = k} \|D^\alpha f\|^p_{L^p(\Omega)} \right)^{1/p}, & \text{if } 1 \le p < \infty, \\ \max_{|\alpha| = k} \|D^\alpha f\|_{L^\infty(\Omega)}, & \text{if } p = \infty. \end{cases}$$

The Sobolev space $W^{k,p}(\Omega)$ is a Banach space with the Sobolev norm. When $p = 2$, it is a Hilbert space and is denoted by $H^k(\Omega) = W^{k,2}(\Omega)$. The corresponding norm and semi–norm are given by $\|\cdot\|_k$ and $|\cdot|_k$, respectively. For $1 \le p < \infty$, note that $C^\infty(\Omega)$ is dense in $W^{k,p}(\Omega)$, and the closure of $C_0^\infty(\Omega)$ in $W^{k,p}(\Omega)$ is denoted by

$W_0^{k,p}(\Omega)$. Like above, $H_0^1(\Omega)$ implies $W_0^{1,2}(\Omega)$.

It is possible to define the Sobolev space $W^{k,p}(\Omega)$ for non–integer k. For $k = m+s$ where m is a non–negative integer and $0 < s < 1$, the Sobolev norm is given by

$$\|f\|_{k,p} = \begin{cases} \left(\|f\|_{m,p}^p + \displaystyle\sum_{|\alpha|=m} \int_\Omega \int_\Omega \frac{|D^\alpha f(x) - D^\alpha f(y)|^p}{|x-y|^{d+sp}}\, dx\, dy \right)^{1/p}, & \text{if } 1 \leq p < \infty, \\ \max \left\{ \|f\|_{m,\infty},\ \displaystyle\max_{|\alpha|=m} \operatorname*{ess\,sup}_{\substack{x,y\in\Omega \\ x\neq y}} \frac{|D^\alpha f(x) - D^\alpha f(y)|}{|x-y|^s} \right\}, & \text{if } p = \infty. \end{cases}$$

We now introduce Sobolev spaces $W^{k,p}(\Omega)$ for $k < 0$. Let $< \cdot, \cdot >$ be the duality pairing defined between $W_0^{-k,q}(\Omega)$ and the set of linear functionals on it, i.e., for a linear functional f

$$< f, g >= f(g), \quad \text{for all } g \in W_0^{-k,q}(\Omega).$$

The Sobolev space $W^{k,p}(\Omega)$ is defined as the dual space of $W_0^{-k,q}(\Omega)$, where $\frac{1}{p} + \frac{1}{q} = 1$. The corresponding norm is given by

$$\|f\|_{k,p} = \sup_{\substack{\phi \in W_0^{-k,q}(\Omega) \\ \phi \neq 0}} \frac{< f, \phi >}{\|\phi\|_{-k,q}}.$$

2. The Sobolev embedding theorem

In this subsection, we present the Sobolev embedding theorem which describes some inclusion relations between Sobolev spaces. For example, the inclusion $L^p(\Omega) \subset L^q(\Omega)$ is continuous, i.e.,

$$\|u\|_{L^q(\Omega)} \leq C\|u\|_{L^p(\Omega)}, \quad \text{for all } u \in L^p(\Omega),$$

for $1 \leq q \leq p \leq \infty$ when Ω is an open bounded set. This is a simple consequence of Hölder's inequality. Here and in the remainder of this dissertation, C is a generic positive constant which depends on d, p, q and Ω. If needed, the dependence of C will be mentioned explicitly.

Theorem A.2 (Theorem 1.4.3.2 in [28]). *Let Ω be a bounded and open set in $\mathbb{R}^d$ with a Lipschitz boundary. If $0 \leq t < s$, then the inclusion of $W^{s,p}(\Omega) \subset W^{t,p}(\Omega)$ is compact for $1 < p < \infty$.*

Theorem A.3 (Sobolev Embedding Theorem [19, 28]). *Let Ω be a bounded and open set in $\mathbb{R}^d$ with a Lipschitz boundary. For $t \leq s$ and $p \leq q < \infty$ such that $s - d/p = t - d/q$,the following inclusion holds:*

$$W^{s,p}(\Omega) \subset W^{t,q}(\Omega).$$

In nonlinear analysis in later chapters, it is sometimes important to control the maximum-norm. We introduce the Sobolev's inequality which represent the Sobolev embedding theorem when q is infinity.

Theorem A.4 (Sobolev's Inequality [13]). *Let Ω be bounded and open in $\mathbb{R}^d$ with a Lipschitz boundary. Let $s > 0$ and $1 \leq p < \infty$ such that*

$$s \leq d \text{ when } p = 1,$$
$$s > d/p \text{ when } p > 1.$$

Then there exists a constant C such that

$$\|u\|_{L^\infty(\Omega)} \leq C\|u\|_{s,p}, \quad \text{for all } u \in W^{s,p}(\Omega).$$

We finish this subsection with the Poincaré inequality involving functions which vanish on the boundary of Ω.

Proposition A.5 (Poincaré Inequality). *Let Ω be a bounded and open set in $\mathbb{R}^d$. For all $u \in W_0^{1,p}(\Omega)$ and $1 \le p < \infty$, there exists a constant C such that*

$$\|u\|_{0,p} \le C|u|_{1,p}.$$

The Poincaré inequality implies that the norm in $W_0^{1,p}(\Omega)$ is equivalent to the semi–norm, i.e., there are two constants C_1 and C_2 such that

$$C_1\|u\|_{1,p} \le |u|_{1,p} \le C_2\|u\|_{1,p}, \quad \text{for all } u \in W_0^{1,p}(\Omega).$$

B. Finite element spaces

In this section, we shall introduce finite element spaces (see, e.g., [13, 19, 30]). We shall give several properties involving operators, such as the Lagrange interpolation operator, the L^2 projection and elliptic projections, mapping from a Sobolev space to a finite space. Finally, inverse and discrete Sobolev inequalities will be given in this section.

For simplicity, let Ω be a bounded open polygon (or polyhedron) in $\mathbb{R}^d$ (for the curved boundary case, see, e.g., [19, 30]). Let $\mathcal{T}_h = \{K\}$ be the triangulation of Ω. Here K is a triangle if $d = 2$ or a tetrahedron if $d = 3$. The mesh size is defined by $h = \max_{K \in \mathcal{T}_h} h_K$, where h_K is the longest side of K. We shall assume that the triangulation is quasi–uniform, i.e., there are two positive constants, β_1 and β_2 independent of h such that for all $K \in \mathcal{T}_h$

(i) $\dfrac{h_K}{h} \ge \beta_1$,

(ii) $\dfrac{\rho_K}{h_K} \ge \beta_2$,

where ρ_K is the diameter of the largest circle inscribed in K.

In this dissertation, we shall only consider finite element spaces contained in

$C^0(\Omega)$. Let r be an integer greater than 1 and $P_r(K)$ be the set of all polynomials with the degree less than r defined on K. The finite element space V_h is given by

$$V_h = \{u \in C^0(\Omega) \mid \ u|_K \in P_r(K) \text{ and } u = 0 \text{ on } \partial\Omega\}.$$

Then V_h is a subspace of $W_0^{1,p}(\Omega)$ for $1 \leq p \leq \infty$. For example, if $r = 1$, then V_h is the piecewise linear finite element space.

Remark B.1. In general, the finite element space V_h is a subspace of $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ for $\alpha < 1/2$ but not for $\alpha > 1/2$. However, it is possible to construct a finite element space which is a subspace of $H^2(\Omega) \cap H_0^1(\Omega)$ (see, e.g., [19, 30]).

Let $\{N_i\}_{i=1}^N$ and $\{\phi_i\}_{i=1}^N$ be the degrees of freedom and the basis for V_h, respectively, where N is the dimension of V_h. We define the Lagrange interpolation operator $I_h \ : \ C^0(\bar{\Omega}) \to V_h$ such that for a given $u \in C^0(\bar{\Omega})$

$$I_h u = \sum_{i=1}^{N} u(N_i)\phi_i.$$

Then we have the following error estimates for I_h:

Theorem B.2. *For $d/2 < s \leq r+1$ and $u \in H^s(\Omega) \cap H_0^1(\Omega)$, there exists a constant C independent of h such that*

(i) $\|u - I_h u\|_0 \leq C h^s |u|_s$,

(ii) $|u - I_h u|_1 \leq C h^{s-1} |u|_s$.

We now introduce the L^2 projection $Q_h \ : \ L^2(\Omega) \to V_h$. Given $u \in L^2(\Omega)$, $Q_h u$ in V_h is the unique function satisfying

$$(Q_h u, v) = (u, v), \quad \text{for all } v \in V_h.$$

Here $(\cdot,\cdot)$ is the L^2–inner–product on Ω. It is easy to see that $\|Q_h u\|_0 \le \|u\|_0$ and

$$\|u - Q_h u\|_0 = \inf_{\chi \in V_h} \|u - \chi\|_0.$$

By Theorem B.2 and interpolation, we have

$$\|u - Q_h u\|_0 \le C h^\alpha |u|_\alpha, \tag{B.1}$$

for $0 \le \alpha \le r+1$. The elliptic projection $P_h : H_0^1(\Omega) \to V_h$ is defined as follows: Given $u \in H_0^1(\Omega)$, $P_h u$ in V_h is the unique function satisfying

$$(P_h u, v)_1 = (u, v)_1, \quad \text{for all } v \in V_h. \tag{B.2}$$

Here $(\cdot,\cdot)_1$ is the H^1 inner product, i.e., for $u,\ v \in H_0^1(\Omega)$

$$(u, v)_1 = \int_\Omega uv\, dx + \int_\Omega \boldsymbol{\nabla} u \cdot \boldsymbol{\nabla} v\, dx. \tag{B.3}$$

The projection P_h satisfies $\|P_h u\|_1 \le \|u\|_1$ and

$$\|u - P_h u\|_1 = \inf_{\chi \in V_h} \|u - \chi\|_1.$$

Furthermore, for $0 \le \alpha \le r$

$$\|u - P_h u\|_1 \le C h^\alpha |u|_{\alpha+1}. \tag{B.4}$$

Note that the L^2 projection and the elliptic projection are orthogonal projectors, and so $Q_h^2 = Q_h$ and $P_h^2 = P_h$.

Finally, we mention some inverse inequalities (see, e.g., [13, 19, 26]) and discrete Sobolev inequalities (see, e.g., [6, 12, 14, 40]).

Theorem B.3 (Inverse Inequalities). *Let $V_h|_K$ be the restriction of V_h onto an element K of $\mathcal{T}_h$. For $1 \le p,\ q \le \infty$ and $0 \le l \le k$, there exists a constant C*

independent of h such that

$$\|u\|_{k,p,K} \le Ch^{l-k+d/p-d/q}\|u\|_{l,q,K},$$

where $\|\cdot\|_{l,s,K}$ denotes the norm in $W^{l,s}(K)$.

Theorem B.4 (Discrete Sobolev Inequalities [6, 13]). *Let the triangulation $\mathcal{T}_h$ be quasi–uniform and let h be the mesh size of the triangulation. For all $u \in V_h$ there is a constant C independent of h such that*

$$\begin{cases} \|u\|_{L^\infty(\Omega)} \le C(1+|\log h|)^{\frac{1}{2}}\|u\|_1 \text{ when } d=2, \\ \|u\|_{L^\infty(\Omega)} \le Ch^{-\frac{1}{2}}\|u\|_1 \text{ when } d=3. \end{cases}$$

C. Technical lemmas

In this section, we shall present a few technical lemmas which will be used often in this dissertation. The first two lemmas bound the integrals of a product of several functions by Sobolev norms. The others deal with the operators, Q_h and P_h.

Lemma C.1. *For $p > d$, $g \in W^{1,p}(\Omega)$, $u \in H^\alpha(\Omega)$ and $v \in H_0^{1-\alpha}(\Omega)$,*

$$\int_\Omega guv_{x_i}\,dx \le C(\Omega,\alpha,p)\|g\|_{1,p}\|u\|_\alpha\|v\|_{1-\alpha},$$

where $v_{x_i} \in H^{-\alpha}(\Omega)$ denotes the partial derivative of v with respect to x_i.

Proof. We have

$$\int_\Omega guv_{x_i}\,dx = \int_\Omega u(gv_{x_i})\,dx \le \|u\|_\alpha\|gv_{x_i}\|_{-\alpha}.$$

To get the desired bound for $\|gv_{x_i}\|_{-\alpha}$, we use interpolation between $H^{-1}(\Omega)$ and $L^2(\Omega)$ (see, e.g., Appendix A in [11], or [27, 28]). By Theorem A.4,

$$\|gv_{x_i}\|_0 \le \|g\|_{L^\infty(\Omega)}\|v_{x_i}\|_0 \le C\|g\|_{1,p}\|v\|_1. \tag{C.1}$$

We shall show that

$$\|gv_{x_i}\|_{-1} \le C\|g\|_{1,p}\|v\|_0 \tag{C.2}$$

and the lemma will follow by interpolation.

Let v be in $C_0^\infty(\Omega)$. Then

$$\begin{aligned}\|gv_{x_i}\|_{-1} &= \sup_{w\in C_0^\infty(\Omega)} \frac{< v, (gw)_{x_i} >}{\|w\|_1} \\ &\le C\|v\|_0 \sup_{w\in C_0^\infty(\Omega)} \frac{\|(gw)_{x_i}\|_0}{\|w\|_1}.\end{aligned}$$

Using Hölder's inequality, we have

$$\begin{aligned}\|g_{x_i}w + gw_{x_i}\|_0 &\le \|g_{x_i}w\|_0 + \|gw_{x_i}\|_0 \\ &\le \|g_{x_i}\|_{L^p(\Omega)}\|w\|_{L^{2q}(\Omega)} + \|g\|_{L^\infty(\Omega)}\|w_{x_i}\|_0,\end{aligned}$$

where $\dfrac{2}{p} + \dfrac{1}{q} = 1$. Applying Theorem A.3, $\|w\|_{L^{2q}(\Omega)} \le C\|w\|_1$, gives

$$\|(gw)_{x_i}\| \le C\|g\|_{1,p}\|w\|_1.$$

This completes the proof of the lemma. □

The above lemma will be used mainly in Chapter IV and the following lemma in Chapter V.

Lemma C.2. *Let p be greater than d. For $k \in L^\infty(\Omega)$, $u \in W_0^{1,p}(\Omega)$, $w \in H_0^1(\Omega)$ and $\varphi \in H_0^1(\Omega)$, we have*

$$\int_\Omega (\nabla u \cdot \nabla w)k\varphi\, dx \le C\|k\|_{L^\infty(\Omega)}\|u\|_{1,p}\|w\|_1\|\varphi\|_1. \tag{C.3}$$

Proof. For $1/p + 1/q = 1$ and $k \in L^\infty(\Omega)$, applying Hölder's inequality gives

$$\int_\Omega (\nabla u \cdot \nabla w)k\varphi\, dx \le C\|k\|_{L^\infty(\Omega)}\|\nabla u\|_{L^p(\Omega)}\|\nabla w\ \varphi\|_{L^q(\Omega)}.$$

Using Hölder's inequality again for $1/r + 1/s = 1$,

$$\|\boldsymbol{\nabla} w\ \varphi\|_{L^q(\Omega)} \leq \|\boldsymbol{\nabla} w\|_{L^{qr}(\Omega)} \|\varphi\|_{L^{qs}(\Omega)}.$$

Taking $r = 2/q$ and $s = 2/(2-q)$, by Theorem A.3:

$$\|\varphi\|_{L^{qs}} \leq C\|\varphi\|_1.$$

This completes the proof of the lemma. $\square$

In our analysis, we need a projection onto V_h which is stable in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$. We have the following proposition:

Proposition C.3 (Appendix in [9]). *Suppose the triangulation $\mathcal{T}_h$ is quasi–uniform. Let $0 \leq \alpha < 1/2$ and Π_h be one of the operators, Q_h or P_h. For all $u \in H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$, there is a constant C independent of h such that*

$$\|\Pi_h u\|_{1+\alpha} \leq C\|u\|_{1+\alpha},$$

i.e., the operator Π_h is stable in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$.

To prove the proposition, we introduce the following lemma for the operators, Q_h and P_h. We skip the proof of the lemma.

Lemma C.4. *Suppose the triangulation $\mathcal{T}_h$ is quasi–uniform. For $0 \leq \alpha < 3/2$,*

(i) $\|u - Q_h u\|_\alpha \leq C\|u\|_\alpha$,

(ii) $\|u - Q_h u\|_1 \leq C h^\alpha \|u\|_{1+\alpha}$.

(iii) $\|u - P_h u\|_1 \leq C h^\alpha \|u\|_{1+\alpha}$.

Here C is independent of h.

Proof of Proposition C.3. See Appendix [9]. $\square$

D. The PCG method

In this section, we shall introduce the PCG method. Its convergence rate also will be presented (see, e.g., [11, 26, 32, 38]). In subsequent sections, the PCG will be used to compute the correction in the inexact Newton method.

Let $A : V_h \to V_h$ be a linear symmetric and positive definite (SPD) operator with respect to an inner product $(\cdot,\cdot)$ on V_h. We shall consider the following linear problem: For a given $f \in V_h$, find $u \in V_h$ such that

$$Au = f. \tag{D.1}$$

Definition D.1. Let λ_{max} and λ_{min} be the largest and smallest eigenvalues of A, respectively. The condition number of A is defined by

$$\kappa(A) = \frac{\lambda_{max}}{\lambda_{min}}.$$

The conjugate gradient (CG) method gives a convergence rate for the problem (D.1) which can be bounded in terms of the condition number of A. To improve the convergence rate of the CG method, we introduce preconditioning.

Specifically, let $B : V_h \to V_h$ be another SPD operator. Then we solve the following problem: For a given $f \in V_h$, find $u \in V_h$ such that

$$BAu = Bf. \tag{D.2}$$

Here we introduce the PCG method based on the inner product $(B^{-1}\cdot,\cdot)$. This inner product makes sense because B^{-1} is also SPD. The algorithm behavior is nothing more than CG applied to (D.2).

Algorithm D.2 (The PCG algorithm). *Let u^0 be an initial iterate. Let $r^0 = f - Au^0$ and $p^0 = z^0 = Br^0$. The sequence of iterates $\{u^k\} \subset V_h$ is generated by* *(i)* $u^{k+1} = u^k + \alpha_k p^k$, *where* $\alpha_k = \dfrac{(r^k, z^k)}{(Ap^k, p^k)}$,

(ii) $r^{k+1} = r^n - \alpha_k A p^k$,

(iii) $z^{k+1} = Br^{k+1}$, $p^{k+1} = z^{k+1} + \beta_k p^k$, *where* $\beta_k = \dfrac{(r^{k+1}, z^{k+1})}{(r^k, z^k)}$.

Let $\kappa = \kappa(BA)$ be the condition number of BA. Since A is SPD, we can define an "energy" norm,

$$\|u\|_A = (Au, u)^{1/2}, \quad \text{for all } u \in V_h.$$

The next theorem bounds the rate of convergence of the PCG method.

Theorem D.3. *Let u be the exact solution to (D.1) and $\{u^k\}$ be the sequence generated in Algorithm D.2. Then*

$$\|u - u^k\|_A \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k \|u - u^0\|_A.$$

CHAPTER III

INEXACT NEWTON METHODS

In this chapter, we shall introduce the inexact Newton method (see, e.g., [1, 22, 24, 35, 42]) which defines an approximate solution to a nonlinear problem, and an abstract theorem which can be used to analyze its convergence will be given.

The Newton method is a classical method used to solve nonlinear problems. However, if the number of unknowns is large then it is often too expensive to find exact Newton corrections. The inexact Newton method is similar to the Newton method except for the way the corrections are defined. In the inexact Newton method the corrections are obtained by defining approximate solutions to the Jacobian systems by an iterative method, such as the PCG method.

A. The inexact Newton method

In this section, we shall consider an abstract nonlinear problem and define the inexact Newton method. Also we shall present the modified inexact Newton method which will be used to define an approximate solution to a nonlinear problem considered in this dissertation.

Our abstract nonlinear problem is defined in terms of two Banach spaces $\mathcal{V}$ and $\mathcal{W}$ (with norms $\|\cdot\|_{\mathcal{V}}$ and $\|\cdot\|_{\mathcal{W}}$). For a bounded linear operator B from $\mathcal{V}$ into $\mathcal{W}$, let $\|\cdot\|_{[\mathcal{V},\mathcal{W}]}$ denote the operator norm,

$$\|B\|_{[\mathcal{V},\mathcal{W}]} = \sup_{\substack{v\in\mathcal{V}\\ v\neq 0}} \frac{\|Bv\|_{\mathcal{W}}}{\|v\|_{\mathcal{V}}}.$$

Let $\mathcal{F}$ be a continuous function mapping $\mathcal{V}$ into $\mathcal{W}'$, the dual space of $\mathcal{W}$. We consider the problem: Find $u \in \mathcal{V}$ satisfying

$$\mathcal{F}(u) = 0. \tag{A.1}$$

Let $\mathcal{F}'(u)$ denote the Fréchet derivative of $\mathcal{F}$ at u.

First of all, we introduce the original inexact Newton method which is developed in [22].

Algorithm A.1 (Inexact Newton Method). *Given an initial iterate $u^0 \in \mathcal{V}$, the sequence of iterates $\{u^k\} \subset \mathcal{V}$ is generated by*

$$u^{k+1} = u^k + \hat{s}^k$$

where $\hat{s}^k$ approximately solves

$$\mathcal{F}'(u^k)s^k = -\mathcal{F}(u^k). \tag{A.2}$$

Specifically, we assume that $\hat{s}^k$ satisfies

$$|||s^k - \hat{s}^k||| \leq \beta \, ||| \, s^k||| \tag{A.3}$$

for some fixed β in $[0, 1)$. Here $||| \cdot |||$ is a norm on $\mathcal{V}$ which is equivalent to $\| \cdot \|_{\mathcal{V}}$.

Let γ_1 and γ_2 be the constants in the norm equivalence relations between $||| \cdot |||$ and $\| \cdot \|_{\mathcal{V}}$, that is,

$$\gamma_1 \; ||| \, v||| \leq \|v\|_{\mathcal{V}} \leq \gamma_2 \; ||| \, v \, ||| \, . \tag{A.4}$$

Remark A.2. Note that the algorithm requires an iterative scheme which gives rise to a reduction in a norm $||| \cdot |||$ which is equivalent to $\| \cdot \|_{\mathcal{V}}$. In general, it is not possible to construct fixed step iterative methods which are convergent in arbitrary norms. Indeed, in almost all of the literature on iterative convergence, convergence reductions are achieved in the L^2 or energy norms. We shall further address this issue in Chapter IV.

We next present the algorithm which we shall consider in this dissertation. It is sometimes referred to as the modified inexact Newton method because $\mathcal{F}^{'}(u^0)$ is used instead of $\mathcal{F}^{'}(u^k)$ in (A.2).

Algorithm A.3 (Modified Inexact Newton Method). *Given an initial iterate $u^0 \in \mathcal{V}$, the sequence of iterates $\{u^k\} \subset \mathcal{V}$ is generated by*

$$u^{k+1} = u^k + \hat{s}^k$$

where $\hat{s}^k$ approximately solves

$$\mathcal{F}'(u^0)s^k = -\mathcal{F}(u^k). \tag{A.5}$$

Specifically, we assume that $\hat{s}^k$ satisfies

$$|||s^k - \hat{s}^k||| \leq \beta \, ||| \, s^k||| \tag{A.6}$$

for some fixed β in $[0, 1)$.

B. A Convergence theorem for the modified inexact Newton method

In this section, the hypotheses under which the sequence generated by Algorithm A.3 converges will be given. The corresponding convergence theorem will be presented.

Minor variations of the algorithms in the previous section have been proposed and studied, for example, in [14, 22, 24, 35]. Our analysis is also a slight modification of theirs.

We consider the following hypotheses:

(H.1) $\mathcal{F}(u) = 0$ has a solution u^* in $\mathcal{V}$.

(H.2) For given $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ and M such that for $\|u - u^*\|_{\mathcal{V}} < \delta$ the Fréchet derivative $\mathcal{F}'(u)$ exists and satisfies:

(H.2.1) $\|\mathcal{F}(u) - \mathcal{F}(u^*) - \mathcal{F}'(u^*)(u - u^*)\|_{\mathcal{W}'} \le \varepsilon \|u - u^*\|_{\mathcal{V}}$

(H.2.2) $\|\mathcal{F}'(u) - \mathcal{F}'(u^*)\|_{[\mathcal{V},\mathcal{W}']} \le \varepsilon$.

(H.2.3) $\mathcal{F}'(u)^{-1}$ exists and satisfies $\|\mathcal{F}'(u)^{-1}\|_{[\mathcal{W}',\mathcal{V}]} \le M$.

The above conditions imply the following theorem (compare to Theorem 2.1 in [14] or Theorem 2.3 in [22]).

Theorem B.1. *Assume that (H.1) and (H.2) hold. Let $t \in (\beta, 1)$ be given where β satisfies (A.6). Then there exists $\delta > 0$ such that if $\|u^0 - u^*\|_{\mathcal{V}} < \delta$, then the sequence of iterates $\{u^k\}$ generated by Algorithm A.3 converges to u^*, in fact,*

$$|||u^{k+1} - u^*||| \le t \, ||| \, u^k - u^* \, ||| \, . \tag{B.1}$$

Remark B.2. The results of our theorem differ from those of [14] in that we only require that the initial iterate is close in a natural norm ($\| \cdot \|_{\mathcal{V}}$) which, in our subsequent applications, is independent of the mesh size. Indeed, although Theorem 5.2 of [14] gives a convergence rate independent of the mesh size, the initial iterate has to be close in a mesh dependent norm.

Remark B.3. The above theorem can be used to guarantee convergence rates independent of the mesh size in PDE applications provided that the functions $\delta(\varepsilon)$, M and bounds for the constants of norm equivalence between $||| \cdot |||$ and $\| \cdot \|_{\mathcal{V}}$ can all be chosen independently of h.

Proof of Theorem B.1. The proof follows the general steps of similar proofs found in the literature. For completeness, we provide a proof that matches the assumptions and setting of Algorithm A.3. It clearly suffices to verify (B.1). We start by observing that

$$\begin{aligned} s^k &= -(\mathcal{F}'(u^0))^{-1}\mathcal{F}'(u^*)(u^k - u^*) \\ &\quad - (\mathcal{F}'(u^0))^{-1}(\mathcal{F}(u^k) - \mathcal{F}(u^*) - \mathcal{F}'(u^*)(u^k - u^*)). \end{aligned} \tag{B.2}$$

Thus,

$$\begin{aligned} |||u^{k+1} - u^*||| = |||(I - (\mathcal{F}'(u^0))^{-1}\mathcal{F}'(u^*))(u^k - u^*) + \hat{s}^k - s^k \\ - (\mathcal{F}'(u^0))^{-1}(\mathcal{F}(u^k) - \mathcal{F}(u^*) - \mathcal{F}'(u^*)(u^k - u^*))\ |||\ . \end{aligned}$$

Let $e_k = u^k - u^*$ and δ be such that (H.2) holds for a positive ε to be determined later. Then,

$$\begin{aligned} |||(I - (\mathcal{F}'(u^0))^{-1}\mathcal{F}'(u^*))e_k||| &\le \gamma_1^{-1}\ \|(\mathcal{F}'(u^0))^{-1}\|_{[\mathcal{W}',\mathcal{V}]}\|\mathcal{F}'(u^0) - \mathcal{F}'(u^*)\|_{[\mathcal{V},\mathcal{W}']}\|e_k\|_{\mathcal{V}} \\ &\le \frac{\gamma_2}{\gamma_1}M\ \varepsilon\ |||\ e_k\ |||\ . \end{aligned}$$

Similarly,

$$|||(\mathcal{F}'(u^0))^{-1}(\mathcal{F}(u^k) - \mathcal{F}(u^*) - \mathcal{F}'(u^*)e_k)||| \le \frac{\gamma_2}{\gamma_1}M\varepsilon\ \ |||\ e_k\ |||\ .$$

Using (A.6) and (B.2) gives

$$\begin{aligned} |||\hat{s}^k - s^k||| &\le \beta\ |||\ s^k||| \\ &\le \beta\{|||(\mathcal{F}'(u^0))^{-1}(\mathcal{F}(u^k) - \mathcal{F}(u^*) - \mathcal{F}'(u^*)e_k)||| \\ &\qquad + |||(\mathcal{F}'(u^0))^{-1}\mathcal{F}'(u^*)e_k|||\} \\ &\le \frac{\gamma_2}{\gamma_1}M\beta\varepsilon\ |||\ e_k\ |||\ +\beta\ |||\ (\mathcal{F}'(u^0))^{-1}\mathcal{F}'(u^*)e_k\ |||\ . \end{aligned}$$

Finally,

$$\begin{aligned} \beta\ |||\ (\mathcal{F}'(u^0))^{-1}\mathcal{F}'(u^*)e_k||| &\le \beta\{|||e_k\ |||\ + |||\ (I - (\mathcal{F}'(u^0))^{-1}\mathcal{F}'(u^*))e_k|||\} \\ &\le (\beta + \frac{\gamma_2}{\gamma_1}M\beta\varepsilon)\ |||\ e_k\ |||\ . \end{aligned}$$

Combining the above inequalities gives

$$|||u^{k+1} - u^*||| \le \left(\beta + \frac{\gamma_2}{\gamma_1} M(3+\beta)\,\varepsilon\right) |||\, u^k - u^*|||$$

and the theorem follows taking $\varepsilon \le \frac{\gamma_1}{\gamma_2 M(3+\beta)}(t-\beta)$. □

It is interesting to note that the continuity constants associated to $\mathcal{F}$ and $\mathcal{F}'$ do not come into the proof. This allows Brown *et al.* [14] to analyze a discrete problem using the L^2 norm on the discrete space $\mathcal{W}$ even though the discrete Fréchet derivatives are not uniformly bounded into this space. The conditions (H.2.1) and (H.2.2) nevertheless hold because the problem considered there only involves linear higher order terms. This fails for our more general application so we are forced to use weaker (negative norm) spaces.

CHAPTER IV

MESH INDEPENDENT CONVERGENCE RESULTS IN $H^{1+\alpha} \cap H_0^1(\Omega)$

The main purposes of this chapter are to develop an iterative method which reduces the error in a discrete norm equivalent to the norm in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ and to obtain the mesh independent convergence rate for the modified inexact Newton method used to compute an approximate solution to the problem (NP).

In this chapter, we shall consider the model nonlinear partial differential equation (NP) which has nonlinearities in the higher order derivatives. We restrict our consideration to problems only when the domain Ω is in $\mathbb{R}^2$ in this chapter. First, we shall set up the problem in a Hilbert space and prove the existence of the solution to the corresponding discrete problem. To define an approximate solution to the problem, the modified inexact Newton method introduced in the previous chapter will be used. Next, an iterative method which defines Newton corrections satisfying (A.6) in Chapter III and the mesh independent convergence rate of the method will be demonstrated. Finally, the results of numerical experiments will be given to support the theory.

A. Hilbert space setting of (NP)

We shall ultimately apply Theorem B.1 in Chapter III to finite element approximations of (NP). Because of the higher order nonlinearity, the hypotheses (H.2.1) and (H.2.2) in Chapter III cannot hold unless functions in $\mathcal{W}$ have two less Sobolev derivatives than those in $\mathcal{V}$. It is common to use the spaces $\mathcal{V} = W^{1,p}(\Omega)$ for $p > 2$ (see e.g., [13, 17]) for the finite element convergence analysis of (NP). Under certain hypotheses on the nonlinearities, it is possible to prove (H.1)-(H.2) using these spaces. To the best of our knowledge, efficient fixed step iterative methods which are convergent in any norm which is equivalent (independently of the discretization parameter) to the norm in $W^{1,p}(\Omega)$ are not known. To get around this issue, we shall analyze our discrete problem in the scale of Sobolev norms in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ for $0 < \alpha < 1/2$. To do this, we start with the analysis of the continuous problem. Most of this analysis will involve proving the inequalities of (H.2) on the continuous level.

To set up the problem, we let φ be in $C_0^\infty(\Omega)$. Then we have

$$(k(u,x)\boldsymbol{\nabla} u, \boldsymbol{\nabla}\varphi) + (\boldsymbol{c} \cdot \boldsymbol{\nabla} u, \varphi) = (f, \varphi)$$

and consider

$$< F(u), \varphi >= (k(u,x)\boldsymbol{\nabla} u, \boldsymbol{\nabla}\varphi) + (\boldsymbol{c} \cdot \boldsymbol{\nabla} u, \varphi) - (f, \varphi). \tag{A.1}$$

To keep the notation from becoming too cumbersome, we have dropped the explicit dependence of u on x above.

We shall use the notation $\boldsymbol{\nabla} k$ to denote the gradient with respect to the x variable considering u independently of x. We shall assume that the quantities

$$k,\ \boldsymbol{\nabla} k,\ \frac{\partial k}{\partial u},\ \frac{\boldsymbol{\nabla}\partial k}{\partial u},\ \frac{\partial^2 k}{\partial^2 u},\ \frac{\boldsymbol{\nabla}\partial^2 k}{\partial^2 u},\ \frac{\boldsymbol{\nabla}\partial k}{\partial x_i},\ \text{and}\ \ \frac{\partial^3 k}{\partial^3 u}$$

are all uniformly bounded independently of $u \in H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$.

Definition A.1. For a given $\alpha \in (0,\ \frac{1}{2})$, we set $V = H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$, $W = H_0^{1-\alpha}(\Omega)$ and $W' = H^{\alpha-1}(\Omega)$.

Note that we have denoted the spaces V and W in contrast to the $\mathcal{V}$ and $\mathcal{W}$ that will be their discrete counterparts, i.e., $\mathcal{V} \subset V$ and $\mathcal{W} \subset W$. The latter pair $(\mathcal{V},\ \mathcal{W})$ takes part in the actual inexact Newton iteration used, in practice, to compute the discrete solution.

1. The F mapping

In this subsection, we shall check whether the function F defined in (A.1) is well–defined.

Lemma A.2. *Suppose that f is in W', then $F(u)$ given by (A.1) is a well–defined map of V into W'.*

To prove the above lemma, we shall use Lemma C.1 in Chapter II. We will also fix p in the interval $(2, 2/(1-\alpha)]$ so that the following two Sobolev inequalities (see Theorem A.4 and A.3 in Chapter II) hold:

$$\|w\|_{L^\infty(\Omega)} \le C\|w\|_{1,p}, \quad \text{for all } w \in W^{1,p}(\Omega) \tag{A.2}$$

and

$$\|w\|_{1,p} \le C\|w\|_{1+\alpha}, \quad \text{for all } w \in H^{1+\alpha}(\Omega). \tag{A.3}$$

We shall also use the Sobolev inequality

$$\|w\|_{L^q(\Omega)} \le C\|w\|_{1-\alpha}, \quad \text{for all } w \in H^{1-\alpha}(\Omega) \tag{A.4}$$

which holds provided that $q \le 2/\alpha$.

Proof of Lemma A.2. Let u be in V. It suffices to show that

$$< F(u), \varphi > \le C(u)\|\varphi\|_{1-\alpha} \quad \text{for all } \varphi \in W.$$

Applying Lemma C.1 in Chapter II gives

$$\begin{aligned} < F(u), \varphi > &= \int_\Omega k(u,x)(\boldsymbol{\nabla} u \cdot \boldsymbol{\nabla}\varphi)\,dx + \int_\Omega (\boldsymbol{c}\cdot\boldsymbol{\nabla} u)\varphi\,dx - \int_\Omega f\varphi\,dx \\ &\le C\{\|k(u,x)\|_{1,p}\|u\|_{1+\alpha} + \|u\|_1 + \|f\|_{-1+\alpha}\}\|\varphi\|_{1-\alpha}. \end{aligned}$$

We note that in the distributional sense,

$$\frac{\partial k(u(x),x)}{\partial x_i} = k_u(u,x)u_{x_i} + k_{x_i}(u,x) \tag{A.5}$$

where the second term obviously denotes differentiation of k with respect to x_i (independent of the u dependence on x). Indeed, (A.5) holds for smooth u and easily follows for general $u \in V$ from the density of smooth functions in V. Thus,

$$\|k(u,x)\|_{1,p} \leq C(\|u\|_{1,p} + 1) \leq C(\|u\|_{1+\alpha} + 1). \tag{A.6}$$

This completes the proof of the lemma. □

Remark A.3. From the above discussion, it is clear that $u \in V$ satisfying $F(u) = 0$ provides a weak solution to (NP). The existence and uniqueness of solutions to nonlinear problems is always a delicate issue. In [17], existence and uniqueness of a weak solution to (NP) in $W^{1,p}(\Omega)$ was verified for $p > 2$ in the case when k only depends on u. In this case, our theory will also give a unique solution $u \in V$ (which coincides with that of [17]).

2. The F' mapping

To study the local behavior of F, we need to introduce the Fréchet derivative $F'(u)$ (a linear map from V to W'). As we show in the next proposition, its definition is given by

$$\begin{aligned} < F'(u)w, \varphi >= &\int_\Omega \frac{\partial k(u,x)}{\partial u} w(\boldsymbol{\nabla} u \cdot \boldsymbol{\nabla} \varphi)\, dx \\ &+ \int_\Omega k(u,x)(\boldsymbol{\nabla} w \cdot \boldsymbol{\nabla} \varphi)\, dx + \int_\Omega (\boldsymbol{c} \cdot \boldsymbol{\nabla} w)\varphi\, dx, \end{aligned} \tag{A.7}$$

for all $w \in V$ and $\varphi \in W$. The next proposition also proves (H.2.1) in the continuous case.

Proposition A.4. *For $u \in V$, the Fréchet derivative of F at u is given by (A.7). Moreover, for a given $\delta > 0$, there exists a constant $C = C(\|u\|_{1+\alpha}, \delta)$ such that*

$$\|F(v) - F(u) - F'(u)(v-u)\|_{-1+\alpha} \le C\|v-u\|_{1+\alpha}^2 \tag{A.8}$$

for all v in the ball $B(u,\delta) \equiv \{v \in V \ : \ \|u-v\|_{1+\alpha} < \delta\}$.

Proof. Using the assumptions on k and a similar argument as used in the proof of Lemma A.2, it follows that for $u \in V$, $F'(u)$ given by (A.7) is a linear map of V into W'.

To finish the proof, it suffices to verify (A.8). This is equivalent to showing that for all $\phi \in W$,

$$< F(v) - F(u) - F'(u)(v-u), \phi > \le C\|v-u\|_{1+\alpha}^2 \|\phi\|_{1-\alpha}.$$

A simple computation gives

$$\begin{aligned} &< F(v) - F(u) - F'(u)(v-u), \phi > \\ &\qquad = \int_\Omega \frac{\partial k(u,x)}{\partial u}(v-u)((\nabla v - \nabla u)\cdot \nabla\phi)\, dx \\ &\qquad\quad + \int_\Omega \left(k(v,x) - k(u,x) - \frac{\partial k(u,x)}{\partial u}(v-u)\right)(\nabla v \cdot \nabla \phi)\, dx. \end{aligned} \tag{A.9}$$

By Lemma C.1 in Chapter II, the first integral of the right hand side of (A.9) is bounded by

$$C\left\|\frac{\partial k(u,x)}{\partial u}(v-u)\right\|_{1,p} \|v-u\|_{1+\alpha}\|\phi\|_{1-\alpha}.$$

Using techniques similar to those used in the proof of Lemma A.2 gives

$$\begin{aligned} \left\|\frac{\partial k(u,x)}{\partial u}(v-u)\right\|_{1,p} &\le C\{\|u\|_{1,p} + 1\}\|v-u\|_{1,p} \\ &\le C\|v-u\|_{1+\alpha}. \end{aligned}$$

We next bound the second integral of the right hand side of (A.9). By Lemma C.1

in Chapter II, it suffices to show that

$$\begin{aligned}
&\left\|k(v(x),x)-k(u(x),x)-\frac{\partial k(u,x)}{\partial u}(v(x)-u(x))\right\|_{1,p} \\
&=\left\|\int_u^v \frac{\partial^2 k(s,x)}{\partial s^2}(v(x)-s)\,ds\right\|_{1,p} \\
&\leq C\|v-u\|_{1+\alpha}^2.
\end{aligned} \tag{A.10}$$

We obviously have

$$\begin{aligned}
\left\|\int_u^v \frac{\partial^2 k(s,x)}{\partial s^2}(v(x)-s)\,ds\right\|_{L^p(\Omega)} &\leq C\|v-u\|_{L^\infty(\Omega)}^2 \\
&\leq C\|v-u\|_{1+\alpha}^2.
\end{aligned} \tag{A.11}$$

Finally, we will bound the semi-norm in (A.10). We have

$$\begin{aligned}
&\left\|\nabla\int_u^v \frac{\partial^2 k(s,x)}{\partial s^2}(v(x)-s)\,ds\right\|_{L^p(\Omega)} \\
&\leq \left\|\int_u^v \frac{\nabla\partial^2 k(s,x)}{\partial^2 s}(v(x)-s))\,ds\right\|_{L^p(\Omega)} \\
&+\left\|\nabla v\left(\frac{\partial k(v,x)}{\partial v}-\frac{\partial k(u,x)}{\partial u}\right)-\nabla u\frac{\partial^2 k(u,x)}{\partial u^2}(v-u)\right\|_{L^p(\Omega)}.
\end{aligned}$$

The first term on the right hand side is bounded analogously to (A.11). For the second, we note that

$$\begin{aligned}
&\nabla v(x)\left(\frac{\partial k(v,x)}{\partial v}-\frac{\partial k(u,x)}{\partial u}\right)-\nabla u(x)\frac{\partial^2 k(v,x)}{\partial v^2}(v(x)-u(x)) \\
&=\nabla(v-u)\int_{u(x)}^{v(x)}\frac{\partial k(s,x)}{\partial s}\,ds+\nabla u\int_{u(x)}^{v(x)}\frac{\partial^3 k(s,x)}{\partial s^3}(v(x)-s)\,ds.
\end{aligned}$$

As above we have

$$\left\|\nabla(v-u)\int_{u(x)}^{v(x)}\frac{\partial k(s,x)}{\partial s}\,ds\right\|_{L^p(\Omega)} \leq C\|u-v\|_{1+\alpha}^2$$

and

$$\left\|\nabla u\int_{u(x)}^{v(x)}\frac{\partial^3 k(s,x)}{\partial s^3}(v(x)-s)\,ds\right\|_{L^p(\Omega)} \leq C\|u-v\|_{1+\alpha}^2.$$

This completes the proof of the proposition. □

We can also show that (H.2.2) holds on the continuous level using similar techniques.

Proposition A.5. *Let u be in V. For a given $\delta > 0$, there exists a constant $C = C(\|u\|_{1+\alpha}, \delta)$ such that, for all v in the ball $B(u, \delta)$ in V,*

$$\|F'(v) - F'(u)\|_{[V,W']} \leq C\|v - u\|_{1+\alpha}.$$

Proof. We need to show that for all w in V and φ in W,

$$< (F'(v) - F'(u))w, \varphi > \leq C\|v - u\|_{1+\alpha}\|w\|_{1+\alpha}\|\varphi\|_{1-\alpha}. \tag{A.12}$$

Now

$$\begin{aligned} < (F'(v) - F'(u))w, \varphi > = & \int_\Omega \left(\frac{\partial k(v,x)}{\partial v} - \frac{\partial k(u,x)}{\partial u} \right) (\boldsymbol{\nabla} v \cdot \boldsymbol{\nabla} \varphi) w \, dx \\ & + \int_\Omega \frac{\partial k(u,x)}{\partial u} (\boldsymbol{\nabla}(v - u) \cdot \boldsymbol{\nabla} \varphi) w \, dx \\ & + \int_\Omega (k(v,x) - k(u,x))(\boldsymbol{\nabla} w \cdot \boldsymbol{\nabla} \varphi) \, dx. \end{aligned}$$

The inequality (A.12) can be derived by applying similar techniques as in the proof of Proposition A.4 to the above identity. This completes the proof of the proposition. □

3. Existence and uniqueness of solutions

There are no results available to guarantee the existence of solutions to Problem (NP) in the generality which we have posed it. To proceed with the analysis, we shall need to make the following additional assumptions.

(B.1) (NP) has a solution u^* in V,

(B.2) $F'(u^*)$ satisfies the uniqueness property:

$$< F'(u^*)w, \varphi >= 0 \quad \text{for all } \varphi \in W \text{ implies } w = 0. \tag{A.13}$$

Remark A.6. If we take $k(u(x), x) = k(u)$, $b = 0$, and $\boldsymbol{c}$ to be divergence free, then it is possible to verify the above assumptions. In this case, (B.2) follows from the proof of uniqueness in Theorem 5.1 of [17], and (B.1) follows from the analysis there.

Using (B.2), we will show that $F'(u^*)$ is an isomorphism (see Proposition A.9 below). This fact will be used to verify the existence of a discrete solution in Section 4. To prove the isomorphism property, we need the following two lemmas.

Lemma A.7. *Let* $D(u, \varphi) = \int_\Omega \boldsymbol{\nabla} u \cdot \boldsymbol{\nabla} \varphi \, dx$. *If* u *is in* $H_0^1(\Omega)$ *and satisfies*

$$\sup_{\varphi \in W} \frac{D(u, \varphi)}{\|\varphi\|_{1-\alpha}} < \infty, \tag{A.14}$$

then u *is also in* V. *Furthermore (A.14) provides an equivalent norm to* $\|u\|_{1+\alpha}$ *on* V.

Proof. Let u satisfy the above conditions and define the functional f in W' by

$$< f, \varphi >= D(u, \varphi) \quad \text{for all } \varphi \in W.$$

Clearly, u is the solution to the Dirichlet problem, $u \in H_0^1(\Omega)$ satisfying

$$D(u, \varphi) =< f, \varphi > \quad \text{for all } \varphi \in H_0^1(\Omega). \tag{A.15}$$

Elliptic regularity for (A.15) implies that $u \in H^{1+\alpha}(\Omega)$ and satisfies

$$\|u\|_{1+\alpha} \leq C\|f\|_{-1+\alpha}.$$

This shows that $\|u\|_{1+\alpha}$ is bounded by a multiple of the supremum in (A.14). The bound in other direction follows from Lemma C.1 in Chapter II. $\square$

Lemma A.8. *For a fixed $u \in V$, the map $w \to k(u,x)w$ is an isomorphism from V onto itself.*

Proof. We first show that for $w \in V$, $k(u,x)w$ is also in V. Since both w and u are in V, estimates similar to those in (A.6) give

$$\|k(u,x)w\|_1 \leq C(1 + \|u\|_{1+\alpha})\|w\|_{1+\alpha}$$

from which it immediately follows that $k(u,x)w \in H_0^1(\Omega)$. Finally, by Lemma C.1 in Chapter II,

$$\begin{aligned} D(k(u,x)w,\varphi) &= \int_\Omega \boldsymbol{\nabla}(k(u,x)w) \cdot \boldsymbol{\nabla}\varphi \, dx \\ &= \int_\Omega \left[\frac{\partial k(u,x)}{\partial u} w(\boldsymbol{\nabla} u \cdot \boldsymbol{\nabla}\varphi) + w(\boldsymbol{\nabla} k(u,x) \cdot \boldsymbol{\nabla}\varphi) \right. \\ &\qquad \left. + \, k(u,x)(\boldsymbol{\nabla} w \cdot \boldsymbol{\nabla}\varphi) \right] dx \\ &\leq C\|u\|_{1+\alpha}\|w\|_{1+\alpha}\|\varphi\|_{1-\alpha}. \end{aligned}$$

Applying Lemma A.7 shows that $k(u,x)w$ is in V and satisfies

$$\|k(u,x)w\|_{1+\alpha} \leq C\|w\|_{1+\alpha}.$$

We note that $k(u,x)^{-1}$ satisfies the same assumptions as $k(u,x)$ so that boundedness of the inverse map follows by the same reasoning. This completes the proof of the lemma. $\square$

The next result shows that $F'(u^*)$ is an isomorphism.

Proposition A.9. *$F'(u^*) : V \to W'$ is an isomorphism, i.e., $F'(u^*)^{-1}$ exists, and there exists a positive constant M such that $\|F'(u^*)^{-1}\|_{[W',V]} \leq M$.*

Proof. In this proof, we adapt the idea in the proof of Theorem 5.2 in [17].

Let $T : W' \to V$ be the solution operator for

$$\int_\Omega \boldsymbol{\nabla}(k(u^*,x)w) \cdot \boldsymbol{\nabla}\varphi\, dx =< f, \varphi > \quad \text{for all } \varphi \in W,$$

that is, $Tf = w$. Let $u \in V$ solve (A.15). Since the map $f \to u$ is an isomorphism from W' onto V, Lemma A.7 implies that T is also.

We define an operator $A_1 : V \to W'$ by

$$< A_1 w, \varphi >= -\int_\Omega (\boldsymbol{\nabla} k(u^*,x) \cdot \boldsymbol{\nabla}\varphi) w\, dx + \int_\Omega (\boldsymbol{c} \cdot \boldsymbol{\nabla} w)\varphi\, dx \quad \text{for all } \varphi \in W.$$

Then, for all $\varphi \in W$,

$$\begin{aligned}
\int_\Omega \boldsymbol{\nabla}(k(u^*,x)TF'(u^*)w) \cdot \boldsymbol{\nabla}\varphi\, dx &=< F'(u^*)w, \varphi > \\
&= \int_\Omega \boldsymbol{\nabla}(k(u^*,x)w) \cdot \boldsymbol{\nabla}\varphi\, dx + < A_1 w, \varphi > \\
&= \int_\Omega \boldsymbol{\nabla}(k(u^*,x)w) \cdot \boldsymbol{\nabla}\varphi\, dx + \int_\Omega \boldsymbol{\nabla}(k(u^*,x)TA_1 w) \cdot \boldsymbol{\nabla}\varphi\, dx.
\end{aligned}$$

Therefore, for $w \in V$,

$$TF'(u^*)w = w + TA_1 w. \tag{A.16}$$

It suffices to show that $TF'(u^*)$ is an isomorphism of V onto V. By Lemma C.1 in Chapter II,

$$< A_1 w, \varphi >\leq C\{\|w\|_\alpha \|\varphi\|_{1-\alpha} + \|w\|_1 \|\varphi\|_0\} \leq C\|w\|_1 \|\varphi\|_{1-\alpha}.$$

Since $H^{1+\alpha}(\Omega)$ is compactly embedded in $H^1(\Omega)$ (see Theorem A.2 in Chapter II), A_1 is a compact operator from V into W'. Thus, TA_1 is also compact from V into V. Hence the mapping $TF'(u^*)$ is a linear Fredholm operator with index zero (see, e.g., [43]). Since (B.2) implies $F'(u^*)$ is injective, $TF'(u^*)$ is injective and bijective also. Since it is also continuous, it is an isomorphism. This completes the proof of the proposition. $\square$

To be able to apply finite element duality, we shall need regularity for the adjoint problem. We consider the adjoint operator $(F'(u^*))^*$ defined by

$$< (F'(u^*))^* v, \varphi > \equiv < F'(u^*)\varphi, v > .$$

Clearly, this is well defined for $v \in W$ and $\varphi \in V$. The next proposition shows that it is also well defined for $v \in V$ and $\varphi \in W$ and gives rise to an isomorphism.

Proposition A.10. *$(F'(u^*))^* : V \to W'$ is an isomorphism.*

Proof. By definition,

$$\begin{aligned} < (F'(u^*))^* \varphi, w > &= \int_\Omega \frac{\partial k(u^*, x)}{\partial u^*} w (\boldsymbol{\nabla} u^* \cdot \boldsymbol{\nabla} \varphi)\, dx \\ &+ \int_\Omega k(u^*, x)(\boldsymbol{\nabla} w \cdot \boldsymbol{\nabla} \varphi)\, dx + \int_\Omega (\boldsymbol{c} \cdot \boldsymbol{\nabla} w)\varphi\, dx. \end{aligned}$$

Using arguments similar to those above and Lemma C.1 in Chapter II, it is easy to show that $(F'(u^*))^*$ is a well defined linear map of V into W'.

Since $F'(u^*)$ is an isomorphism from V onto W', $(F'(u^*))^*$ is an isomorphism from W onto V' and

$$\|((F'(u^*))^*)^{-1}\|_{[V',W]} = \|(F'(u^*))^{-1}\|_{[W',V]}.$$

Thus, by Proposition A.9,

$$\|\varphi\|_{1-\alpha} \le M \sup_{u \in V} \frac{< F'(u^*)u, \varphi >}{\|u\|_{1+\alpha}}. \tag{A.17}$$

The above inequality implies that $(F'(u^*))^*$ is injective on V.

Define $A_2 : V \to W'$ by

$$\begin{aligned} < A_2 w, \varphi > = &\int_\Omega \frac{\partial k(u^*, x)}{\partial u^*} (\boldsymbol{\nabla} u^* \cdot \boldsymbol{\nabla} w)\varphi\, dx + \int_\Omega (\boldsymbol{c} \cdot \boldsymbol{\nabla} \varphi) w\, dx \\ &- \int_\Omega \frac{\partial k(u^*, x)}{\partial u^*} (\boldsymbol{\nabla} u^* \cdot \boldsymbol{\nabla} \varphi) w\, dx - \int_\Omega (\boldsymbol{\nabla} k(u^*, x) \cdot \boldsymbol{\nabla} \varphi) w\, dx. \end{aligned} \tag{A.18}$$

Then we have

$$< (F'(u^*))^* w, \varphi >= \int_\Omega \boldsymbol{\nabla}(k(u^*, x)w) \cdot \boldsymbol{\nabla}\varphi \, dx + < A_2 w, \varphi > .$$

Thus, as in the proof of Proposition A.9,

$$T(F'(u^*))^* w = w + TA_2 w$$

and it suffices to show that TA_2 is compact on V. This, in turn, will follow if we show that

$$< A_2 w, \varphi > \leq C\|w\|_{1+\beta}\|\varphi\|_{1-\alpha}. \tag{A.19}$$

for some β with $0 < \beta < \alpha$.

By Lemma C.1 in Chapter II, the last three terms of (A.18) can be bounded by the right hand side of (A.19). For example, the third term is bounded by

$$C\left\| w\frac{\partial k(u^*, x)}{\partial u^*}\right\|_{1,q} \|u^*\|_{1+\alpha}\|\varphi\|_{1-\alpha} \leq C\|w\|_{1+\beta}\|\varphi\|_{1-\alpha}$$

provided that q is taken so that

$$H^{1+\beta}(\Omega) \subset W^{1,q}(\Omega).$$

For the first term in (A.18), we choose below $r = 1/(1-\alpha)$ and $s = 1/\alpha$. Then

$$\begin{aligned}
\int_\Omega \frac{\partial k(u^*, x)}{\partial u^*}(\boldsymbol{\nabla}u^* \cdot \boldsymbol{\nabla}w)\varphi \, dx &\leq \left\|\frac{\partial k(u^*, x)}{\partial u^*}\varphi\boldsymbol{\nabla}u^*\right\|_0 \|\boldsymbol{\nabla}w\|_0 \\
&\leq C\|w\|_1\|\boldsymbol{\nabla}u^*\|_{L^{2r}(\Omega)}\|\varphi\|_{L^{2s}(\Omega)} \\
&\leq C\|w\|_1\|u^*\|_{1,2r}\|\varphi\|_{L^{2s}(\Omega)} \\
&\leq C\|w\|_{1+\beta}\|\varphi\|_{1-\alpha}.
\end{aligned}$$

□

B. Existence of a discrete solution

In this section, we define a finite element approximation of (NP). By applying the results of [17], we will conclude the existence of a finite element solution which is close to the solution u^*.

Recall that $\mathcal{T}_h$ and V_h are a triangulation of Ω with mesh size h and the corresponding finite element space, and inverse inequalities are held for finite element functions.

The discrete counterpart of (NP) reads: Find $u_h^* \in V_h$ such that

$$\int_\Omega k(u_h^*, x)(\boldsymbol{\nabla} u_h^* \cdot \boldsymbol{\nabla}\varphi)\, dx + \int_\Omega (\boldsymbol{c} \cdot \boldsymbol{\nabla} u_h^*)\varphi\, dx = \int_\Omega f\varphi\, dx \quad \text{for all } \varphi \in V_h. \tag{B.1}$$

For $u, \varphi \in V$, let $A(u,\varphi) =< F'(u^*)u, \varphi >= \hat{A}(u,\varphi) + \hat{D}(u,\varphi)$, where

$$\begin{cases} \hat{A}(u,\varphi) = \displaystyle\int_\Omega k(u^*, x)(\boldsymbol{\nabla} u \cdot \boldsymbol{\nabla}\varphi)\, dx, \\ \hat{D}(u,\varphi) = \displaystyle\int_\Omega \frac{\partial k(u^*, x)}{\partial u^*}(\boldsymbol{\nabla} u^* \cdot \boldsymbol{\nabla}\varphi) u\, dx + \int_\Omega (\boldsymbol{c} \cdot \boldsymbol{\nabla} u)\varphi\, dx. \end{cases}$$

We will show that the form $A(\cdot,\cdot)$ satisfies a discrete inf-sup condition. To do this, we need the following few lemmas:

Lemma B.1. *There exist two constants $C_1 > 0$ and C_2 such that*

$$C_1\|u\|_1^2 - C_2\|u\|_{1-\alpha}^2 \le A(u,u) \quad \textit{for all } u \in V. \tag{B.2}$$

Proof. By the assumption on k and Poincaré's inequality, there exists $C > 0$ satisfying

$$\hat{A}(u,u) \ge C\|u\|_1^2 \quad \text{for all } u \in H_0^1(\Omega). \tag{B.3}$$

By the Schwarz inequality,

$$\left| \int_\Omega (\boldsymbol{c} \cdot \boldsymbol{\nabla} u) u\, dx \right| \le C\|\boldsymbol{\nabla} u\|_0 \|u\|_0 \le C\|u\|_1 \|u\|_{1-\alpha}. \tag{B.4}$$

For $1/p + 1/q = 1$, applying Hölder's inequality gives

$$\left| \int_\Omega \frac{\partial k(u^*, x)}{\partial u^*} (\boldsymbol{\nabla} u^* \cdot \boldsymbol{\nabla} u) u \, dx \right| \leq C \|\boldsymbol{\nabla} u^*\|_{L^p(\Omega)} \|\boldsymbol{\nabla} u \; u\|_{L^q(\Omega)}. \tag{B.5}$$

Using Hölder's inequality again for $1/r + 1/s = 1$,

$$\|\boldsymbol{\nabla} u \; u\|_{L^q(\Omega)} \leq \|u\|_{L^{qr}(\Omega)} \|\boldsymbol{\nabla} u\|_{L^{qs}(\Omega)}. \tag{B.6}$$

We take $p = 2/(1-\alpha)$, $q = 2/(1+\alpha)$ and $r = (1+\alpha)/\alpha$ and apply the Sobolev inequalities (A.3) and (A.4) to get

$$\left| \int_\Omega \frac{\partial k(u^*, x)}{\partial u} (\boldsymbol{\nabla} u^* \cdot \boldsymbol{\nabla} u) u \, dx \right| \leq C \|u^*\|_{1+\alpha} \|u\|_1 \|u\|_{1-\alpha}. \tag{B.7}$$

Combining $ab \leq (\epsilon a^2)/2 + b^2/(2\epsilon)$, (B.3), (B.4) and (B.7), we get the result. □

Remark B.2. The proof of the above lemma implies that there exists a constant C such that

$$|A(u, v)| \leq C \|u\|_1 \|v\|_1 \quad \text{for all } u, \quad v \in V.$$

Lemma B.3. *For each $u \in V$, there exists $h_0 > 0$ such that for h less than h_0, the problem : Find $u_h \in V_h$ satisfying*

$$A(u_h, \varphi) = A(u, \varphi) \quad \textit{for all } \varphi \in V_h \tag{B.8}$$

has a unique solution which we denote by $T_h u$. Furthermore,

$$\|u - T_h u\|_1 \leq C h^\alpha \|u\|_{1+\alpha}. \tag{B.9}$$

Proof. The proof immediately follows Lemma B.1 and the finite element duality argument (see, e.g., [39]) with Proposition A.10. □

Lemma B.4. *Let h be less than h_0 defined in Lemma B.3. Then*

$$\|T_h u\|_{1+\alpha} \leq C \|u\|_{1+\alpha}$$

.

Proof. Let $e = u - T_h u$ and $Q_h : V \to V_h$ be the $L^2(\Omega)$ projection onto V_h. Then, by Proposition C.3 in Chapter II,

$$\begin{cases} \|u - Q_h u\|_1 \leq C h^{\alpha} \|u\|_{1+\alpha}, \\ \\ \|Q_h u\|_{1+\alpha} \leq C \|u\|_{1+\alpha}. \end{cases} \tag{B.10}$$

By an inverse inequality, (B.10) and (B.9)

$$\begin{aligned} \|e\|_{1+\alpha}^2 &\leq \|T_h u - Q_h u\|_{1+\alpha}^2 + \|Q_h u - u\|_{1+\alpha}^2 \\ &\leq C h^{-2\alpha} \|T_h u - Q_h u\|_1^2 + C \|u\|_{1+\alpha}^2 \\ &\leq C \|u\|_{1+\alpha}^2. \end{aligned}$$

Hence,

$$\|e\|_{1+\alpha} \leq C \|u\|_{1+\alpha}.$$

The triangle inequality used for $T_h u = e + u$ completes the proof of this lemma. □

Proposition B.5 (Discrete Inf-Sup Condition). *Let h be less than h_0 defined in Lemma B.3. Then there exists a constant C such that*

$$\|\varphi\|_{1-\alpha} \leq C \sup_{u_h \in V_h} \frac{A(u_h, \varphi)}{\|u_h\|_{1+\alpha}} \quad \textit{for all } \varphi \in V_h. \tag{B.11}$$

Proof. By (A.17) and Lemma B.4, for all $\varphi \in V_h$

$$\begin{aligned} \|\varphi\|_{1-\alpha} &\leq C \sup_{u \in V} \frac{A(u, \varphi)}{\|u\|_{1+\alpha}} = C \sup_{u \in V} \frac{A(T_h u, \varphi)}{\|u\|_{1+\alpha}} \\ &\leq C \sup_{u \in V} \frac{A(T_h u, \varphi)}{\|T_h u\|_{1+\alpha}} \leq C \sup_{u_h \in V_h} \frac{A(u_h, \varphi)}{\|u_h\|_{1+\alpha}}. \end{aligned}$$

□

Remark B.6. Similar to the derivation of (A.17), (B.11) implies

$$\|u_h\|_{1+\alpha} \leq C \sup_{\varphi \in V_h} \frac{A(u_h, \varphi)}{\|\varphi\|_{1-\alpha}} \quad \text{for all } u_h \in V_h. \tag{B.12}$$

Finally, we are ready to prove the existence of the discrete solution for the model problem by applying a result in [17].

Theorem B.7 (Existence of The Discrete Solution). *With assumptions (B.1) and (B.2), there exist two constants $\delta > 0$ and $h_0 > 0$ such that for $h \leq h_0$ there exists a unique solution u_h^* for problem (B.1) in the ball $B(u^*, \delta)$. Moreover there exists a constant C independent of h such that*

$$\|u^* - u_h^*\|_{1+\alpha} \leq C \inf_{\xi \in V_h} \|u^* - \xi\|_{1+\alpha}. \tag{B.13}$$

Proof. Proposition A.4 and A.5 show $F'(u)$ exists for all $u \in V$ and is Lipschitz continuous in a neighborhood of u^*. Moreover, $F'(u^*)$ is an isomorphism from V to W' (see Proposition A.9). The theorem follows from the discrete inf-sup condition (B.12) and Theorem 7.1 in [17]. □

Remark B.8. The results of the previous section show that the solution has regularity, $u^* \in H^{1+s}(\Omega) \cap H_0^1(\Omega)$ for $s < 1/2$. Taking $s > \alpha$ and applying the above theorem gives

$$\|u^* - u_h^*\|_{1+\alpha} \leq C h^{s-\alpha} \|u^*\|_{1+s},$$

i.e., u_h^* converges to u^* in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$.

C. The discrete problem in the framework of Chapter III

In this section, we set up the discrete problem in the framework of Chapter III. We start by defining $\mathcal{V} = \mathcal{W} = V_h$ with norms $\|\cdot\|_{1+\alpha}$ on $\mathcal{V}$ and $\|\cdot\|_{1-\alpha}$ on $\mathcal{W}$. We identify $\mathcal{W}'$ with V_h and define for $G \in \mathcal{W}'$,

$$< G, \varphi > \equiv (G, \varphi) \quad \text{for all } \varphi \in \mathcal{W}.$$

We then define $F_h : \mathcal{V} \mapsto \mathcal{W}'$ by

$$(F_h(v), \varphi) = \int_\Omega k(v, x)(\boldsymbol{\nabla} v \cdot \boldsymbol{\nabla}\varphi)\, dx + \int_\Omega (\boldsymbol{c} \cdot \boldsymbol{\nabla} v)\varphi\, dx - \int_\Omega f\varphi\, dx, \tag{C.1}$$

for all $v \in \mathcal{V}$ and $\varphi \in \mathcal{W}$. Clearly, the problem of finding $u_h^* \in \mathcal{V}$ satisfying $F_h(u_h^*) = 0$ coincides with the discrete problem (B.1). Thus, (H.1) is contained in Theorem B.7.

For $u \in V$, we define the linear map $F_h' : \mathcal{V} \to \mathcal{W}'$ by

$$\begin{aligned}(F_h'(u)v, \varphi) = &\int_\Omega k(u, x)(\boldsymbol{\nabla} v \cdot \boldsymbol{\nabla}\varphi)\, dx \\ &+ \int_\Omega \frac{\partial k(u, x)}{\partial u}(\boldsymbol{\nabla} u \cdot \boldsymbol{\nabla}\varphi)v\, dx + \int_\Omega (\boldsymbol{c} \cdot \boldsymbol{\nabla} v)\varphi\, dx,\end{aligned} \tag{C.2}$$

for all $v \in \mathcal{V}$ and $\varphi \in \mathcal{W}$. Note that

$$(F_h'(u)v, \varphi) =< F'(u)v, \varphi >, \quad \text{for all } v \in \mathcal{V},\ \varphi \in \mathcal{W},$$

i.e., $F_h'(u)$ is the restriction of $F'(u)$ to $\mathcal{V} \times \mathcal{W}$. Because of this, Proposition A.4 immediately implies that for $u \in \mathcal{V}$, $\delta > 0$ and $v \in B(u, \delta) \subset \mathcal{V}$,

$$\|F_h(v) - F_h(u) - F_h'(u)(v - u)\|_{\mathcal{W}'} \leq C(\|u\|_{1+\alpha}, \delta)\|v - u\|_{\mathcal{V}}^2.$$

Thus we verified (H.2.1). Similarly, by Proposition A.5, for u and v as above,

$$\|F_h'(v) - F_h'(u)\|_{[\mathcal{V},\mathcal{W}']} \leq C(\|u\|_{1+\alpha}, \delta)\|v - u\|_{\mathcal{V}} \tag{C.3}$$

and (H.2.2) follows directly.

Let g be in $\mathcal{W}'$ and extend g to a functional on W by

$$< g, \varphi > \equiv (g, \varphi) \quad \text{for all } \varphi \in W.$$

Then

$$\|g\|_{W'} = \sup_{\varphi \in W} \frac{(g, \varphi)}{\|\varphi\|_{1-\alpha}}$$
$$\leq C \sup_{\varphi \in W} \frac{(g, Q_h\varphi)}{\|Q_h\varphi\|_{1-\alpha}} = C\|g\|_{\mathcal{W}'}.$$

We used the fact that the L^2 projection Q_h is a bounded operator on $\| \cdot \|_{1-\alpha}$. Let $u = (F'(u^*))^{-1}g$. Using the fact that $F'(u^*)$ is an isomorphism and (B.12) gives that the solution $u_h \in V_h$ of

$$A(u_h, \theta) = A(u, \theta) = (g, \theta) \quad \text{for all } \theta \in V_h$$

satisfies

$$\|u_h\|_{\mathcal{V}} \leq M\|g\|_{\mathcal{W}'}, \tag{C.4}$$

i.e., $\|(F_h'(u^*))^{-1}\|_{[\mathcal{W}',\mathcal{V}]} \leq M$. Here M can be chosen independent of h if h_0 is small enough.

The final condition (H.2.3) required for the application of the results of Chapter III is contained in the following proposition.

Proposition C.1. *There exist $h_0 > 0$ and $\delta > 0$ such that if h is less than h_0 and u_h is in the ball $B(u_h^*, \delta)$ in $\mathcal{V}$, $F_h'(u_h)^{-1} : \mathcal{W}' \to \mathcal{V}$ exists and satisfies*

$$\|F_h'(u_h)^{-1}\|_{[\mathcal{W}',\mathcal{V}]} \leq 2M. \tag{C.5}$$

Proof. By (C.3) and (C.4), there exists δ_0 such that

$$\|I - (F_h'(u^*))^{-1}F_h'(u)\|_{[\mathcal{V},\mathcal{V}]} \leq \|(F_h'(u^*))^{-1}\|_{[\mathcal{W}',\mathcal{V}]}\|F_h'(u^*) - F_h'(u)\|_{[\mathcal{V},\mathcal{W}']}$$
$$\leq CM\|u^* - u\|_{1+\alpha},$$

for all u in $B(u^*, \delta_0) \subset V$. If we choose $2\delta < \min\{\frac{1}{2CM}, \delta_0\}$, then

$$\|I - (F_h'(u^*))^{-1}F_h'(u)\|_{[\mathcal{V},\mathcal{V}]} < \frac{1}{2}, \quad \text{for all } u \in B(u^*, 2\delta) \subset V.$$

By the Neumann series argument, $F_h'(u)$ is nonsingular and $\|F_h'(u)^{-1}\|_{[\mathcal{W}',\mathcal{V}]} \leq 2M$.

By Remark B.8, we can choose h_0 sufficiently small so that

$$\|u^* - u_h^*\|_{1+\alpha} \leq \delta$$

when h is less than h_0. Then for $u_h \in B(u_h^*, \delta) \subset \mathcal{V}$,

$$\|u_h - u^*\|_{1+\alpha} \leq 2\delta$$

and the conclusion of the proposition follows. □

The above results show that (H.1) and (H.2) hold for our discrete framework. Moreover, the functions $\delta(\varepsilon)$ and M can be chosen independent of h if h_0 is small enough. Thus, the modified inexact Newton of the form given in Chapter III will converge at a uniform rate (independently of mesh size h) if an iteration satisfying (A.6) in Chapter III is developed (see Section D).

Remark C.2. The above proposition shows that $F_h'(u_h)$ is an isomorphism if $u_h \in V_h$ is close enough to u_h^*, i.e., there are two constants C_1 and C_2 independent of h such that

$$C_1\|\varphi\|_{1+\alpha} \leq \|F_h'(u_h)\varphi\|_{-1+\alpha} \leq C_2\|\varphi\|_{1+\alpha} \quad \text{for all } \varphi \in V_h. \tag{C.6}$$

D. An iteration satisfying (A.6) in Chapter III

In this section, we define an iteration which satisfies (A.6) in Chapter III when $\mathcal{V}$ is defined as in the previous section. We start by defining computable Sobolev norms by using a variation of the approach from [7, 10]. An iteration satisfying (A.6) in Chapter III is then constructed in terms of these norms.

We assume that the space V_h results from a multilevel sequence of meshes. Specifically, we assume that we have a sequence of nested triangulations, e.g., the triangles in $\mathcal{T}_{j+1}$ are formed by subdividing those in $\mathcal{T}_j$ into four by connecting the midpoints of the edges. We require that $\mathcal{T}_1$ is of unit size and set V_j to be the finite element space corresponding to $\mathcal{T}_j$. Let h_j be the mesh size of V_j, for $j = 1, \ldots, J$, and let $h = h_J$. Then $V_h \equiv V_J$.

We next define a sequence of approximation operators $\widehat{Q}_j : L^2(\Omega) \to V_j$. Let $\{\phi_i^{(j)}\}_{i=1}^{m_j}$ be the nodal basis for V_j. For $j > 0$, set

$$\widehat{Q}_j u = \sum_{i=1}^{m_j} \frac{(u, \phi_i^{(j)})}{(1, \phi_i^{(j)})} \phi_i^{(j)}. \tag{D.1}$$

Define

$$T_s u = \sum_{j=1}^{J} h_j^{-2s} \widehat{Q}_j u \quad \text{for all } u \in V_h. \tag{D.2}$$

Then, there are two constants $C_0 = C_0(s)$ and $C_1 = C_1(s)$ not depending on h such that for $-3/2 < s < 0$,

$$C_0 \|u\|_s \leq (T_s u, u)^{1/2} \leq C_1 \|u\|_s \quad \text{for all } u \in V_h. \tag{D.3}$$

We note that if we set $\widehat{Q}_0 = 0$ and define T_s by

$$T_s u = \sum_{j=1}^{J} h_j^{-2s} (\widehat{Q}_j - \widehat{Q}_{j-1})^2 u \quad \text{for all } u \in V_h, \tag{D.4}$$

then (D.3) still holds (see, e.g., [7]).

Let $\widehat{F} = F_h'(u^0)$ where $u^0 \in V_h$ is the starting iterate of Algorithm A.3 in Chapter III and satisfies $\|u^0 - u_h^*\|_{1+\alpha} \leq \delta$ so that (H.1) and (H.2) hold. In addition, define $a(\cdot,\cdot)$ on $V_h \times V_h$ and $A : V_h \to V_h$ by

$$\begin{cases} a(u, v) = (T_{-1+\alpha} \widehat{F} u, \widehat{F} v), \\ (Au, v) = a(u, v) \quad \text{for all } u, v \in V_h. \end{cases} \tag{D.5}$$

Then $a(\cdot,\cdot)$ is clearly symmetric and positive definite by Proposition C.1. We define a discrete norm $|||\cdot|||$ on V_h by

$$|||u_h||| \equiv a(u_h, u_h)^{1/2} \quad \text{for all } u_h \in V_h.$$

It follows from (C.6) and (D.3) that $|||\cdot|||$ is uniformly (independently of h) equivalent to $\|\cdot\|_{1+\alpha}$ on V_h. Recall that this equivalence was required by (A.4) in Chapter III.

We then have the following lemma.

Lemma D.1. *There exist two positive constants $\tilde{C}_0$ and $\tilde{C}_1$ independent of h such that*

$$\tilde{C}_0 a(u,u) \le a(T_{-1-\alpha}Au, u) \le \tilde{C}_1 a(u,u) \quad \textit{for all } u \in V_h. \tag{D.6}$$

Proof. For all u in V_h, using Remark C.2 gives

$$\begin{aligned} a(T_{-1-\alpha}Au, u) &= (T_{-1-\alpha}Au, Au) \ge C\|Au\|^2_{-1-\alpha} \\ &= C \sup_{v\in V} \frac{(Au,v)^2}{\|v\|^2_{1+\alpha}} \ge C\frac{a(u,u)^2}{\|u\|^2_{1+\alpha}} \\ &\ge \tilde{C}_0 a(u,u). \end{aligned}$$

On the other hand, by (D.3)

$$\begin{aligned} a(T_{-1-\alpha}Au, u) &\le C \sup_{v\in V} \frac{(Au,v)^2}{\|v\|^2_{1+\alpha}} \\ &= C \sup_{v\in V} \frac{(Au,Q_h v)^2}{\|v\|^2_{1+\alpha}}. \end{aligned}$$

Using the boundedness of Q_h on V gives

$$\begin{aligned} a(T_{-1-\alpha}Au, u) &\le C \sup_{v_h\in V_h} \frac{a(u,v_h)^2}{\|v_h\|^2_{1+\alpha}} \\ &\le \tilde{C}_1 a(u,u). \end{aligned}$$

This completes the proof of the lemma. $\square$

We consider the following problem: Find s^k in V_h satisfying

$$a(s^k, \varphi) = (-T_{-1+\alpha} F_h(u^k), \widehat{F}\varphi) \quad \text{for all } \varphi \in V_h. \tag{D.7}$$

The solution to (D.7) and (A.2) in Algorithm A.3 in Chapter III coincide. To define $\hat{s}^k$, we apply the m step PCG method (see Section D in Chapter II) to (D.7) with the zero initial iterate. The preconditioner used here is $T_{-1-\alpha}$. Then

$$|||s^k - \hat{s}^k||| \le \frac{2q^m}{1+q^{2m}} \; |||\, s^k|||, \tag{D.8}$$

where $q = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} < 1$ and $\kappa = cond(T_{-1-\alpha}A) \le \tilde{\kappa} \equiv \tilde{C}_1/\tilde{C}_0$. Thus, (A.6) in Chapter III holds for

$$\beta = \frac{2\tilde{q}^m}{1+\tilde{q}^{2m}} < 1$$

where $\tilde{q} = \frac{\sqrt{\tilde{\kappa}}-1}{\sqrt{\tilde{\kappa}}+1} < 1$ and independent of h.

Remark D.2. We can apply the one step PCG (which is the steepest descent) method also. In this case, (A.6) in Chapter III holds for $\beta = \dfrac{\tilde{\kappa}-1}{\tilde{\kappa}+1} < 1$.

We can now conclude with the following theorem.

Theorem D.3. *Suppose that we use the iterative method described above for computing the approximation $\hat{s}^k$. There is a positive number h_0 and a $\delta > 0$ such that if $h \le h_0$ and $\|u^0 - u_h^*\|_{1+\alpha} \le \delta$, then the modified inexact Newton algorithm converges monotonically with a geometric rate of convergence which is independent of h.*

E. Numerical results

In this section, numerical results supporting Theorem D.3 for a model problem will be given. We shall illustrate uniform and linear convergence rates of the modified inexact Newton method. Finally, the efficiency of the method will be described comparing with the Newton method. Here and in the remainder of this dissertation, all numerical results are computed using C/C++ code under the system, Pentium III with 548.636 MHz CPU and 256M RAM.

We shall present the results for (A.6) in Chapter III defined using the algorithm of the previous section for the two cases when T_s is given by (D.2) and (D.4) applied to the following problem:

$$\begin{aligned} -\mathrm{div}(k(u,x)\boldsymbol{\nabla}u) + \boldsymbol{c}\cdot\boldsymbol{\nabla}u = f, \quad x \in \Omega \\ u(x) = 0, \quad x \in \partial\Omega. \end{aligned} \tag{E.1}$$

Here, $\Omega = (0,1)\times(0,1)$, $k(u,x) = 1/(u^2+1) + e^{-x_1^2-x_2^2}$, $\boldsymbol{c} = (1,1)$, and the exact solution is $u^* = u(x_1,x_2) = (x_1 - x_1^2)(x_2 - x_2^2)$. The right hand side f is defined by applying the left hand side to the exact solution.

The discrete problem is obtained by using linear basis functions on triangles of mesh size $h = 1/2^n$, $n = 5,6,7,8,9$. To define $\hat{s}^k$ in the modified inexact Newton method, we use the five step PCG method for (D.7), i.e., $m = 5$ in (D.8). The norm (or Sobolev space) we use is corresponded to $\alpha = 0.05$. We stop the algorithm when the normalized discrete l_2 norm of the nodal values of $F_h(u_h^k)$ is less than 10^{-6}. The initial nonlinear iterate u_h^0 of the method is $0.9 * I_h u^*$ in Table E.1 and E.2 where $I_h u^*$ is the interpolant of the exact solution.

Table E.1 and E.2 illustrate the number of nonlinear iterations required to reach the above-mentioned convergence criteria. We also report the L^2 and H^1 norm errors between the discrete solution and $I_h u^*$. The number of nonlinear iterations increase slightly as h decreases. The rate of increase decreases for smaller h. Similar behavior is observed in [7] when T_{-1} is applied as a preconditioner for the Laplacian, an iterative procedure which rate of convergence can also be bounded independently of the number of unknowns. The operator $T_{-1-\alpha}A$ is better conditioned with the choice of T_s given by (D.4) and this is in agreement with the fact that the results of Table E.1 are better than ones of Table E.2.

Table E.1. Nonlinear iteration numbers, $T_s = \sum h_j^{-2s}(\widehat{Q}_j - \widehat{Q}_{j-1})^2$

h^{-1}	nonlinear iterations	$\|u_h - I_h u^*\|_0$	$\lvert u_h - I_h u^*\rvert_1$
32	12	5.50e-05	2.50e-04
64	14	1.38e-05	6.29e-05
128	16	3.49e-06	1.59e-05
256	18	9.00e-07	4.19e-06
512	19	2.61e-07	1.87e-06

Table E.2. Nonlinear iteration numbers, $T_s = \sum h_j^{-2s}\widehat{Q}_j$

h^{-1}	nonlinear iterations	$\|u_h - I_h u^*\|_0$	$\lvert u_h - I_h u^*\rvert_1$
32	22	5.50e-05	2.50e-04
64	28	1.38e-05	6.27e-05
128	33	3.48e-06	1.58e-05
256	37	9.03e-07	4.37e-06
512	39	3.04e-07	2.54e-06

To illustrate the linear convergence in Theorem D.3, we set up a problem where the exact discrete solution was known. To do this, we applied the discrete nonlinear operator to $I_h u^*$ so that $I_h u^*$ was the exact discrete solution. Figure 1 shows the linear convergence with respect to the norm $|||\cdot|||$ when $h = 1/64$ and $T_s = \sum h_j^{-2s}(\widehat{Q}_j - \widehat{Q}_{j-1})^2$. For these results, we took again $m = 5$ steps of PCG to define $\hat{s}^k$.

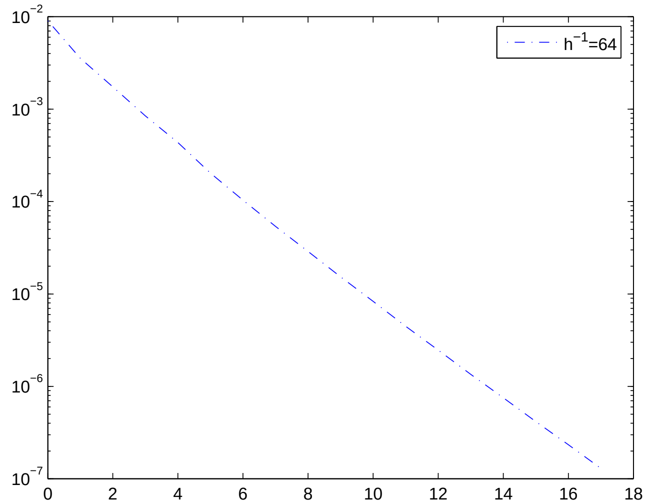

Fig. 1. Linear convergence with five steps of PCG in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$

Figure 2 illustrates that the algorithm converges even if only *one step* of the PCG is taken in the definition of $\hat{s}^k$ (Remark D.2). The results correspond to a problem with $h = 1/64$ and $T_s = \sum h_j^{-2s}(\widehat{Q}_j - \widehat{Q}_{j-1})^2$. One step of PCG results in $\hat{s}^k$ being a fairly crude approximation to s^k which leads to a very slow convergence for the nonlinear problem. Nevertheless, Figure 2 illustrates the monotone convergence behavior guaranteed in Theorem D.3.

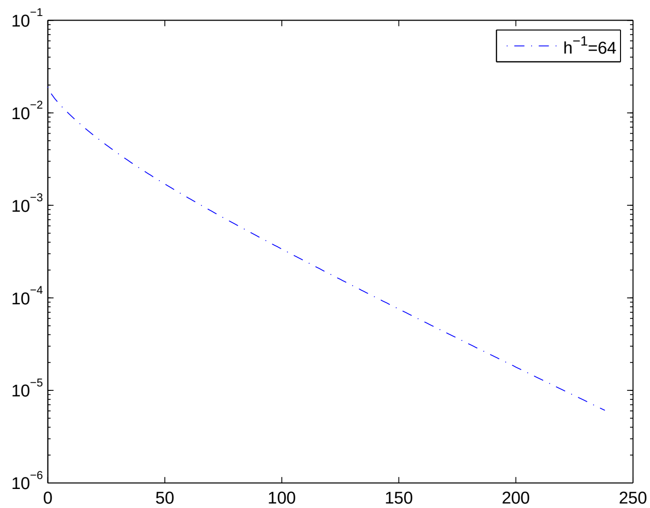

Fig. 2. Linear convergence with one step of PCG in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$

Next, we present in Table E.3 the dependence of the nonlinear iteration number on the accuracy in (D.8) given by the (PCG) iteration number m, for the test problem with $h = 1/64$. If the number of linear iterations (PCG iteration number) increases then the number of nonlinear iterations decreases because β in Algorithm A.3 in Chapter III is getting smaller. One notices the significant difference between one and two PCG iterations.

Table E.3. Various PCG steps in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$

PCG steps	nonlinear iterations	CPU time
1	281	45.9460s
2	41	8.5467s
3	32	8.1308s
4	21	6.2960s
5	14	4.8063s

Table E.4 and E.5 compare the total linear (inner) iteration numbers and the elapsed CPU times between the modified inexact and exact Newton methods when $u_h^0 = 0.5 * I_h u$. We can see that the modified inexact Newton method is more efficient than the exact Newton method if the number of unknowns is large.

Table E.4. Comparison between inexact Newton and Newton - Inner iteration numbers in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$

h^{-1}	inexact Newton	Newton	ratio (%)
32	70	145	207
64	85	180	218
128	100	219	219
256	105	268	255
512	115	327	284

Table E.5. Comparison between inexact Newton and Newton - Running time in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$

h^{-1}	inexact Newton	Newton	ratio (%)
32	0.8329s	1.1848s	142
64	4.1614s	5.9971s	144
128	20.3479s	30.7013s	151
256	89.6894s	159.0548s	177
512	421.0510s	844.0037s	201

CHAPTER V

MESH INDEPENDENT CONVERGENCE RESULTS IN $H_0^1(\Omega)$

In this chapter, we shall provide convergence rates of modified inexact Newton methods applied to nonlinear second order problems in the Sobolev space $H_0^1(\Omega)$. Specifically, we shall consider the model problem:

$$\begin{aligned} -\operatorname{div}(k(u)\boldsymbol{\nabla} u) + \boldsymbol{c}\cdot\boldsymbol{\nabla} u = f, &\quad x \in \Omega \\ u(x) = 0, &\quad x \in \partial\Omega. \end{aligned} \tag{NP2}$$

Here k and $\boldsymbol{c}$ are smooth functions on $\bar{\Omega}$ which is a bounded polygonal (or polyhedral) domain in $\mathbb{R}^d$, for $d = 2,\ 3$. In addition, k is bounded away from zero, and $\boldsymbol{c}$ is divergence free. We shall assume that the quantities

$$k,\ \frac{\partial k}{\partial u},\ \text{and}\ \frac{\partial^2 k}{\partial^2 u}$$

are all uniformly bounded independently of $u \in V$. To keep the notation from becoming too cumbersome, we have dropped the explicit dependence of u on x in our notation.

In Chapter IV, we obtained convergence rates of the modified inexact Newton method applied to the problem in the scale Hilbert spaces $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$. In this chapter, we shall establish similar results in $H_0^1(\Omega)$.

Recall that in the analysis for nonlinear problems it is important to bound the norm in $L^\infty(\Omega)$. Since the norm in $H_0^1(\Omega)$ does not coerce the norm in $L^\infty(\Omega)$, we shall analyze the problem in Sobolev spaces $W_0^{1,p}(\Omega)$ where p is greater than d. In discrete spaces, we shall bound the norm in $L^\infty(\Omega)$ by the norm in $H_0^1(\Omega)$ using discrete Sobolev inequalities (see Theorem B.4 in Chapter II).

In Chapter IV, one of the main results was to construct an iterative scheme to define Newton corrections satisfying the error reduction (A.6) in Chapter III. Since the corrections defined in this chapter will satisfy (A.6) in a norm which is equivalent to the norm in $H_0^1(\Omega)$, we can use domain decomposition methods (see, e.g., [16]), multigrid methods (see, e.g., [11]), or the iterative scheme presented in Chapter IV, setting $\alpha = 0$.

This chapter is organized as follows. In Section A, we analyze the model problem in Sobolev spaces $W_0^{1,p}(\Omega)$ and show existence and uniqueness of the solution to the corresponding discrete problem. A uniform convergence rate of the modified inexact Newton method in the norm in $H_0^1(\Omega)$ is given in Section B. The numerical experiments to support the convergence theory are presented in the last section. For the sake of convenience, we use the iterative scheme which is developed in Chapter IV to define Newton corrections.

A. Sobolev space setting of (NP2)

In this section, we set up the model problem in Sobolev spaces and present a few propositions related to the operators corresponding to the model problem. These propositions are used to verify the assumptions (H.1) and (H.2) in Chapter III under which the modified inexact Newton method converges. Finally, we define a discrete approximation of the solution to (NP2). Most results in this section are based on *Caloz and Rappaz's* paper, [17].

To set up the problem, we let φ be in $C_0^\infty(\Omega)$. Then we have

$$(k(u)\boldsymbol{\nabla} u, \boldsymbol{\nabla}\varphi) + (\boldsymbol{c} \cdot \boldsymbol{\nabla} u, \varphi) = (f, \varphi).$$

Recall that for $p > d$ the norm in $L^\infty(\Omega)$ is bounded by the norm in $W_0^{1,p}(\Omega)$ by the Sobolev embedding, i.e.,

$$\|u\|_\infty \le C\|u\|_{1,p}, \quad \text{for all } u \in W_0^{1,p}(\Omega). \tag{A.1}$$

In this chapter, unless specified, we assume that p is greater than d and $1/p + 1/q = 1$. Let $V = W_0^{1,p}(\Omega)$, $W = W_0^{1,q}(\Omega)$ and $W' = W^{-1,p}(\Omega)$ be the dual space of W.

The nonlinear operator F corresponding to the problem (NP2) is given by

$$< F(u), \varphi >= (k(u)\boldsymbol{\nabla} u, \boldsymbol{\nabla}\varphi) + (\boldsymbol{c}\cdot\boldsymbol{\nabla} u, \varphi) - (f, \varphi), \tag{A.2}$$

for all $u \in V, \quad \varphi \in W$.

To study the local behavior of F, we need to define the Fréchet derivative $F'(u)$ of F at $u \in V$. It is defined by

$$< F'(u)w, \varphi >= \int_\Omega \boldsymbol{\nabla}(k(u)w)\cdot\boldsymbol{\nabla}\varphi\, dx + \int_\Omega (\boldsymbol{c}\cdot\boldsymbol{\nabla} w)\varphi\, dx, \tag{A.3}$$

where $w \in V$ and $\varphi \in W$.

Remark A.1. For $d = 2$, the authors of [17] showed that F and $F'(u)$ are well defined from V to W'. In the same way, we can prove that they are also well defined for $d = 3$.

We have the following two propositions which verify the first two of the hypothesis (H.2). The fact that the norm in V coerces the norm in $L^\infty(\Omega)$ plays an important role to obtain these propositions.

Proposition A.2. *For a given $\delta > 0$, there exists a constant $C = C(\|u\|_{1,p}, \delta)$ such that*

$$\|F(v) - F(u) - F'(u)(v-u)\|_{-1,p} \le C\|v-u\|_{1,p}^2 \tag{A.4}$$

for all v in the ball $B(u,\delta) \equiv \{v \in V \ : \ \|u-v\|_{1,p} < \delta\}$.

Proof. The proof is very similar to the proof of Proposition A.4 in Chapter IV. For completeness, we provide the proof.

To finish the proof, it suffices to verify that for all $\phi \in W$,

$$< F(v) - F(u) - F'(u)(v-u), \phi > \leq C\|v-u\|_{1,p}^2 \|\phi\|_{1,q}.$$

A simple computation gives

$$\begin{aligned} &< F(v) - F(u) - F'(u)(v-u), \phi > \\ &\quad = \int_\Omega \frac{\partial k(u)}{\partial u}(v-u)((\boldsymbol{\nabla} v - \boldsymbol{\nabla} u) \cdot \boldsymbol{\nabla} \phi)\, dx \\ &\qquad + \int_\Omega \left(k(v) - k(u) - \frac{\partial k(u)}{\partial u}(v-u) \right) (\boldsymbol{\nabla} v \cdot \boldsymbol{\nabla} \phi)\, dx. \end{aligned} \tag{A.5}$$

By Hölder and Sobolev inequalities, we control the first term in the right hand side of (A.5) as follows:

$$\begin{aligned} &\int_\Omega \frac{\partial k(u)}{\partial u}(v-u)((\boldsymbol{\nabla} v - \boldsymbol{\nabla} u) \cdot \boldsymbol{\nabla} \phi)\, dx \\ &\leq C \left\| \frac{\partial k(u)}{\partial u}(v-u) \right\|_{L^\infty(\Omega)} \|v-u\|_{1,p} \|\phi\|_{1,q} \\ &\leq C\|v-u\|_{L^\infty(\Omega)} \|v-u\|_{1,p} \|\phi\|_{1,q} \\ &\leq C\|v-u\|_{1,p}^2 \|\phi\|_{1,q}. \end{aligned}$$

Since

$$\begin{aligned} &\left\| k(v(x)) - k(u(x)) - \frac{\partial k(u)}{\partial u}(v(x) - u(x)) \right\|_{L^\infty(\Omega)} \\ &= \left\| \int_u^v \frac{\partial^2 k(s)}{\partial s^2}(v(x) - s)\, ds \right\|_{L^\infty(\Omega)} \\ &\leq C\|v-u\|_{L^\infty(\Omega)}^2 \\ &\leq C\|v-u\|_{1,p}^2, \end{aligned}$$

the second integral in (A.5) is bounded, i.e.,

$$\begin{aligned}
&\int_\Omega \left(k(v) - k(u) - \frac{\partial k(u)}{\partial u}(v-u)\right)(\nabla v \cdot \nabla \phi)\, dx \\
&\leq C\|v-u\|_{1,p}^2 \|v\|_{1,p}\|\phi\|_{1,q} \\
&\leq C(\|u\|_{1,p},\delta)\|v-u\|_{1,p}^2\|\phi\|_{1,q}.
\end{aligned}$$

Because of $v \in B(u,\delta)$, we have the last inequality. This completes the proof of the proposition. $\square$

Proposition A.3. *Let u be in V. For a given $\delta > 0$, there exists a constant $C = C(\|u\|_{1,p},\delta)$ such that, for all v in the ball $B(u,\delta)$ in V,*

$$\|F'(v) - F'(u)\|_{[V,W']} \leq C\|v-u\|_{1,p}.$$

Proof. Like the identity in the proof of Proposition A.5 in Chapter IV, we have the following identity:

$$\begin{aligned}
< (F'(v) - F'(u))w, \varphi > = &\int_\Omega \left(\frac{\partial k(v)}{\partial v} - \frac{\partial k(u)}{\partial u}\right)(\nabla v \cdot \nabla \varphi) w\, dx \\
&+ \int_\Omega \frac{\partial k(u)}{\partial u}(\nabla (v-u) \cdot \nabla \varphi) w\, dx \\
&+ \int_\Omega (k(v) - k(u))(\nabla w \cdot \nabla \varphi)\, dx.
\end{aligned} \tag{A.6}$$

Using the similar techniques as in the proof of Proposition A.2 and the above identity completes the proof of the proposition. $\square$

Next, we consider the existence of a solution to the model problem. As we mentioned in Chapter IV, there are no results available to guarantee the existence of solutions to Problem (NP2) in general. To proceed with the analysis, we shall need to make the following assumptions.

(S.1) (NP2) has a solution u^* in V,

(S.2) $F'(u^*)$ satisfies the uniqueness property:

$$< F'(u^*)w, \varphi >= 0, \quad \text{for all } \varphi \in H_0^1(\Omega) \text{ implies } w = 0. \tag{A.7}$$

(S.3) (Regularity) For a given $f \in L^p(\Omega)$, consider

$$\begin{cases} -\Delta u = f \text{ in } \Omega, \\ u = 0 \text{ on } \partial\Omega. \end{cases} \tag{A.8}$$

Then there exists a unique solution which satisfies

$$\|u\|_{2,p} \leq C\|f\|_p. \tag{A.9}$$

Remark A.4. In [17], (S.1) and (S.2) were verified when the boundary of Ω is smooth and $d = 2$.

Remark A.5. In two dimension, if Ω is a convex polygon then (S.3) holds when $1 < p \leq 2 + \epsilon$ for a $\epsilon > 0$ (see, e.g., [27]). In three dimension, if Ω is a rectangular parallelepiped then (S.3) is true for $\frac{6}{5} \leq p < \infty$ (see, e.g, [20, 21]).

Remark A.6. In [21], the author presented that the Laplace operator defined in the problem (A.8) is an isomorphism from $W_0^{1,p}(\Omega)$ to $W^{-1,p}(\Omega)$ when $\frac{3}{2} - \epsilon < p \leq 6 - d + \epsilon$ for a $\epsilon > 0$.

Under the above assumptions, we can show that $F'(u^*)$ and the adjoint $(F'(u^*))^*$ which will be defined below are isomorphisms. This facts will be used to prove the existence of the discrete solution.

Proposition A.7. *Let p satisfy the conditions in Remark A.6. Then $F'(u^*) : V \to W'$ is an isomorphism.*

Proof. The proof of this proposition is given in [17] when the boundary of Ω is smooth. In the proof, the authors used the fact that the inverse of the Laplacian operator is an isomorphism from $W^{-1,p}(\Omega)$ to $W_0^{1,p}(\Omega)$ and an compact operator from $L^p(\Omega)$ to $W_0^{1,p}(\Omega)$. Thanks to Remark A.6 and (S.3), we complete the proof. $\square$

Next, we consider the adjoint operator $(F'(u^*))^*$ defined by

$$< (F'(u^*))^* v, \varphi > \equiv < F'(u^*)\varphi, v > .$$

Clearly, this is well defined for $v \in W$ and $\varphi \in V$. The next proposition says that the adjoint operator is also well defined for $v \in V$ and $\varphi \in W$. In addition, it gives rise to an isomorphism.

Proposition A.8. *Let p satisfy the conditions in Remark A.6. Then $(F'(u^*))^* : V \to W'$ is an isomorphism, i.e., $(F'(u^*)^*)^{-1}$ exists and there is a positive constant M such that $\|((F'(u^*))^*)^{-1}\|_{[W',V]} \leq M$.*

Proof. In this proof, we use the same argument as in the proof of Proposition A.10 in Chapter IV. Since $F'(u^*)$ is isomorphism, we have the following inf–sup condition,

$$\|\varphi\|_{1,q} \leq C \sup_{u \in V} \frac{(F'(u^*)u, \varphi)}{\|u\|_{1,p}}.$$

The above inequality implies that $(F'(u^*))^*$ is injective on V.

Let T be the inverse of the minus Laplacian operator. Then $T : W' \to V$ is an isomorphism. We can easily check that for $w \in V$

$$T \frac{(F'(u^*))^*}{k(u^*)} w = w + T\Big(\frac{\boldsymbol{c}}{k(u^*)} \cdot \boldsymbol{\nabla} w\Big). \tag{A.10}$$

By (S.3) and the fact that $W^{2,p}(\Omega)$ is compactly embedded in $W^{1,p}(\Omega)$, T is compact from $L^p(\Omega)$ to V. Then $T\dfrac{(F'(u^*))^*}{k(u^*)}$ is a Fredholm operator with index zero.

Therefore, since $(F'(u^*))^*$ is injective and continuous, $(F'(u^*))^*$ is also isomorphism. $\square$

We now define a finite element approximate solution of (NP2). The discrete counterpart of (NP2) reads: Find $u_h \in V_h$ such that

$$(F(u_h), \varphi) = 0, \quad \text{for all } \varphi \in V_h. \tag{A.11}$$

Then the nonlinear problem has a unique discrete solution.

Theorem A.9 (Existence of the Discrete Solution). *With assumptions (S.1), (S.2) and (S.3), there exist two constants $\delta > 0$ and $h_0 > 0$ such that for $h \leq h_0$ there exists a unique solution u_h^* to the problem (A.11) in the ball $B(u^*, \delta)$. Moreover there exists a constant C independent of h such that*

$$\|u^* - u_h^*\|_{1,p} \leq C \inf_{\xi \in V_h} \|u^* - \xi\|_{1,p}. \tag{A.12}$$

Proof. In [17], the authors proved these results only in $\mathbb{R}^2$. It is not difficult to extend the results in $\mathbb{R}^3$. □

B. Convergence rates in the norm of $H_0^1(\Omega)$

In this section, we obtain a uniform convergence result for the modified inexact Newton method in the norm of $H_0^1(\Omega)$. The main tools of the analysis are discrete Sobolev inequalities,

$$\begin{cases} \|u\|_{L^\infty(\Omega)} \leq C(1 + |\log h|)^{\frac{1}{2}} \|u\|_1 \text{ when } d = 2, \\ \\ \|u\|_{L^\infty(\Omega)} \leq C h^{-\frac{1}{2}} \|u\|_1 \text{ when } d = 3. \end{cases} \tag{B.1}$$

We define a mesh dependent constant,

$$R_h := \begin{cases} 1 + |\log h| \text{ when } d = 2, \\ \\ h^{-1} \text{ when } d = 3. \end{cases} \tag{B.2}$$

We recall that the exact solution u^* is in V.

Under the framework of Chapter III, let $\mathcal{V} = V_h$ and $\mathcal{W} = V_h$ with the norm in $H_0^1(\Omega)$. Let $\mathcal{F} = F_h$ and $\mathcal{F}' = F_h'$, where F_h and F_h' are defined as in Chapter IV.

Let $\|\cdot\|_*$ be a norm in V_h which is equivalent to the norm in $H_0^1(\Omega)$. We assume that this equivalence is independent of the mesh size. Then we have the following convergence theorem for the modified inexact Newton method.

Theorem B.1. *Suppose that we have an iterative method which defines an approximation of Newton corrections satisfying the error reduction* (A.6) *in Chapter III. Then there are positive numbers h_0 and $\delta(h)$, such that if $h \leq h_0$ and $\|u^0 - u_h^*\|_* \leq \delta(h)$ then the modified inexact Newton method converges monotonically with a geometric rate of convergence which is independent of h.*

Remark B.2. If we take $\|\cdot\|_* = |||\cdot|||$ with $\alpha = 0$, the error reduction $\|s^k - \hat{s}^k\|_* \leq \beta\|s^k\|_*$ can be verified with the iterative method introduced in Chapter IV. We will see numerical results for this norm in the next section. Also we can use multigrid methods (see, e.g. [2, 11, 18, 41]) to define Newton corrections.

Remark B.3. Unfortunately δ in Theorem B.1 depends on the mesh size. It means that the initial iterate of the modified inexact Newton method should be chosen closer to the exact solution as the number of unknowns are getting larger.

The proof of the above theorem is same as the proof of Theorem B.1 in Chapter III if the hypotheses (H.1) and (H.2) in Chapter III are valid. In the rest of this section, we shall verify these hypotheses.

Note that the hypothesis (H.1) of the existence of the discrete solution has been verified in Theorem A.9.

Using the discrete Sobolev inequalities (B.1), Proposition A.2, A.3 and Lemma C.2 in Chapter II, the next proposition is obtained. It verifies (H.2.1) and (H.2.2).

Proposition B.4. *For a given $\delta > 0$, there exists a constant $C = C(\|u_h\|_1, \delta)$ such that if $\|v_h - u_h\|_1 < \delta$ then*

(a) $\|F_h(v_h) - F_h(u_h) - F_h'(u_h)(v_h - u_h)\|_{-1} \leq CR_h\|v_h - u_h\|_1^2$,

(b) $\|F_h'(v_h) - F_h'(u_h)\|_{[\mathcal{V},\mathcal{W}']} \leq CR_h\|v_h - u_h\|_1$.

Proof. To finish the proof of (a), it is enough to show that

$$< F_h(v_h) - F_h(u_h) - F_h'(u_h)(v_h - u_h), \varphi > \leq CR_h\|v_h - u_h\|_1^2\|\varphi\|_1,$$

for all $\varphi \in W$.

By (A.5), we have

$$\begin{aligned} &< F_h(v_h) - F_h(u_h) - F_h'(u_h)(v_h - u_h), \varphi > \\ &\leq C\{\|v_h - u_h\|_{L^\infty(\Omega)}\|v_h - u_h\|_1 + \|v_h - u_h\|_{L^\infty(\Omega)}^2\|\varphi\|_1\}. \end{aligned}$$

Applying the discrete Sobolev inequalities completes the proof of (a).

By (A.6) in the proof of Proposition A.3 and the similar arguments as above, (b) is obtained easily. □

Finally, (H.2.3) follows from the next proposition.

Proposition B.5. *If h is less than a small constant $h_0 > 0$ then there exist constants δ dependent on h and M independent of h such that if $\|u_h - u_h^*\|_1 < \delta$ then*

$$\|F_h'(u_h)^{-1}\|_{[\mathcal{W}',\mathcal{V}]} \leq M.$$

Proof. We sketch the proof, step by step.

Step 1: $F'(u^*)$ is an isomorphism from $H_0^1(\Omega)$ to $H^{-1}(\Omega)$:

Using (S.2) and the techniques in the proof of Proposition A.7 verifies Step 1.

Step 2: For a given $\delta_1 > 0$, there exists a constant $C = C(\delta_1, \|u^*\|_{1,p})$ such that if $\|v - u^*\|_{1,p} < \delta_1$ then $\|F'(v) - F'(u^*)\|_{[H_0^1, H^{-1}]} \leq C\|v - u^*\|_{1,p}$:

It is enough to show that

$$< (F'(v) - F'(u^*))w, \varphi > \leq C\|v - u^*\|_{1,p}\|w\|_1\|\varphi\|_1,$$

for all $u,\ \varphi \in H_0^1(\Omega)$. We can prove this using the equations and the arguments in the proof of Proposition A.3 and Lemma C.2 in Chapter II.

Step 3: If h is less than a small constant $h_0 > 0$ then there exists a constant $M_1 > 0$ independent of h such that $\|F'(u_h^*)^{-1}\|_{[H_0^1, H^{-1}]} \leq M_1$:

This step follows from Step 1, 2, (A.12) and the Neumann series argument.

Step 4: There is a mesh–independent constant M_2 such that $\|F_h'(u_h^*)^{-1}\|_{[\mathcal{W}', \mathcal{V}]} \leq M_2$ if h is small enough:

We consider the following problem. For a given $u \in H_0^1(\Omega)$, find $u_h \in V_h$ satisfying that

$$(F_h'(u_h^*)u_h, \phi) =< F'(u_h^*)u, \phi >, \quad \text{for all } \phi \in V_h. \tag{B.3}$$

Using Lemma C.2 in Chapter II, we can show that $F'(u_h^*)$ satisfies a Gärding type inequality, for $1 \leq q < 3/2$,

$$C_1\|u\|_1^2 - C_2\|u\|_{1,q}^2 \leq < F'(u_h^*)u, u >, \quad \text{for all } u \in V,$$

and is bounded, i.e.,

$$| < F'(u_h^*)u, v > | \leq C\|u\|_1\|v\|_1, \quad \text{for all } u, \quad v \in V.$$

Furthermore, since the adjoint operator of $F'(u_h^*)$ is an isomorphism from V to W', if h is small enough then by the duality argument we have $\|u - u_h\|_{1,q} \leq$

$Ch^\beta\|u-u_h\|_1$ for a positive β. By Schatz's argument (see, e.g., [39]), (B.3) has a unique solution u_h and

$$\|u_h\|_1 \le C_s\|u\|_1. \tag{B.4}$$

Hence, using Step 3 and (B.4), we obtain that $F'(u_h^*)$ satisfies a discrete inf–sup condition, i.e., there exists a constant $M_2 = C_sM_1$ satisfying

$$\begin{aligned}
\|\varphi\|_1 &\le M_1 \sup_{u\in V} \frac{(F'(u_h^*)u,\varphi)}{\|u\|_1} \\
&= M_1 \sup_{u\in V} \frac{(F_h'(u_h^*)u_h,\varphi)}{\|u\|_1} \\
&\le C_sM_1 \sup_{u\in V} \frac{(F_h'(u_h^*)u_h,\varphi)}{\|u_h\|_1} \\
&\le C_sM_1 \sup_{u_h\in V_h} \frac{(F_h'(u_h^*)u_h,\varphi)}{\|u_h\|_1}
\end{aligned}$$

for all $\varphi \in V_h$. The following is also true: For all $u_h \in V_h$,

$$\|u_h\|_1 \le M_2 \sup_{\varphi\in V_h} \frac{(F_h'(u_h^*)u_h,\varphi)}{\|\varphi\|_1}.$$

Here M_2 is independent of h.

Since $F'(u_h^*)$ is an isomorphism from $H_0^1(\Omega)$ to $H^{-1}(\Omega)$, for all $g \in V_h$ there exists $u \in H_0^1(\Omega)$ such that $u = F'(u_h^*)^{-1}g$. By (B.3), the solution to the problem

$$(F_h'(u_h^*)u_h,\phi) =< F'(u_h^*)u,\phi >= (g,\phi), \quad \text{for all } \phi \in V_h$$

satisfies $\|u_h\|_{\mathcal{V}} \le M_2\|g\|_{\mathcal{W}'}$, i.e., $\|F_h'(u_h^*)^{-1}\|_{[\mathcal{W}',\mathcal{V}]} \le M_2$.

Step 5: There exist two constants δ and M

$$\|F_h'(u_h)^{-1}\|_{[\mathcal{W}',\mathcal{V}]} \le M, \tag{B.5}$$

if $\|u_h - u_h^*\|_1 < \delta$:

By Proposition B.4, Step 4, and the Neumann series argument, we can choose $\delta \leq \dfrac{1}{2CM_2R_h}$ and $M = 2M_2$ satisfying (B.5). Here M is independent of h.

$\square$

C. Numerical results

In this section, we shall present numerical results for the modified inexact Newton method applied to the problem (E.1) in Chapter IV when $k(u,x) = k(u) = 1/(1+u^2)$. Our experiments will be given only in two dimension. We shall investigate whether the convergence rates are uniform and linear. In addition, the efficiency of the modified inexact Newton method will be described.

In the previous section, we obtained mesh independent convergence rates of the modified inexact Newton method if the distance between each initial iterate and the exact solution is less than CR_h^{-1}. For simplicity, in the numerical experiments, we choose the initial iterates $u_h^0 = 0.95 * I_h u^*$, where $I_h u^*$ is the linear interpolation of u^*. We use the same scheme as in Chapter IV taking $\alpha = 0$. To define the Newton direction, the five step PCG method is applied with the preconditioner T_{-1}.

Table C.1 and C.2 show that the convergence rates are independent of the mesh size. Here we use two different preconditioners, $T_{-1} = \sum h_j^2(\widehat{Q}_j - \widehat{Q}_{j-1})^2$ and $T_{-1} = \sum h_j^2 \widehat{Q}_j$. Like in Chapter IV, the former preconditioner is better than the latter.

Table C.1. Nonlinear iteration numbers, $T_{-1} = \sum h_j^2(\widehat{Q}_j - \widehat{Q}_{j-1})^2$

h^{-1}	nonlinear iterations	$\|u_h - I_h u^*\|_0$	$\|u_h - I_h u^*\|_1$
32	8	5.30e-05	2.42e-04
64	9	1.33e-05	6.07e-05
128	10	3.39e-06	1.54e-05
256	10	9.36e-07	4.38e-06
512	10	3.82e-07	2.43e-06

Table C.2. Nonlinear iteration numbers, $T_{-1} = \sum h_j^2 \widehat{Q}_j$

h^{-1}	nonlinear iterations	$\|u_h - I_h u^*\|_0$	$\|u_h - I_h u^*\|_1$
32	13	5.30e-05	2.42e-04
64	15	1.33e-05	6.09e-05
128	17	3.40e-06	1.56e-05
256	18	9.93e-07	4.67e-06
512	19	3.95e-07	2.52e-06

Figures 3 and 4 describe that the convergence rates are linear when $h^{-1} = 64$. In Figures 3 and 4, five step and one step PCG methods are applied, respectively. Obviously, if we use the one step PCG method, the convergence is slow.

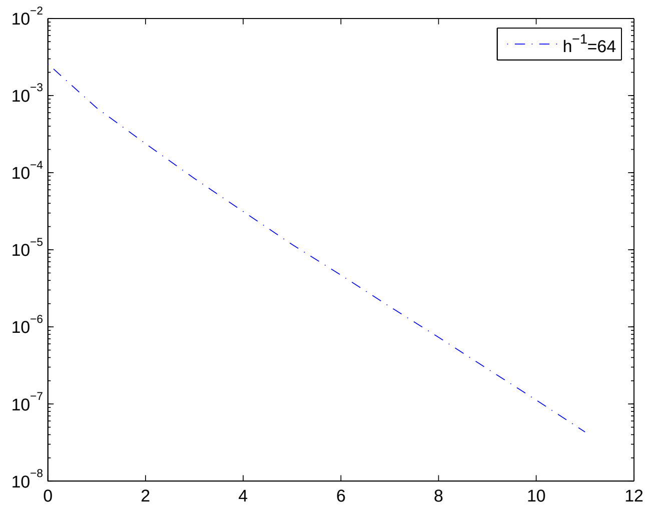

Fig. 3. Linear convergence with five steps of PCG in $H_0^1(\Omega)$

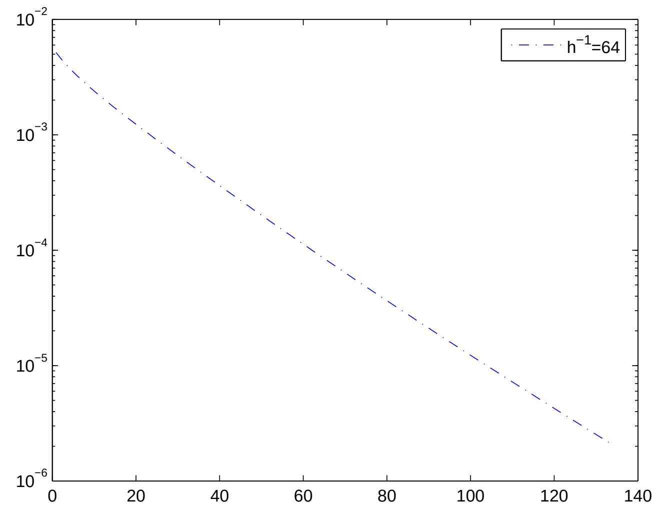

Fig. 4. Linear convergence with one step of PCG in $H_0^1(\Omega)$

Table C.3 presents the dependence of the nonlinear iteration number on the accuracy in (D.8) given by the (PCG) iteration number m. We also present the elapsed CPU time. It is interesting to note that there is a big difference between $m = 1$ and $m = 2$.

Table C.3. Various PCG steps in $H_0^1(\Omega)$

PCG steps	nonlinear iterations	CPU time
1	136	17.2954s
2	21	3.3345s
3	16	3.0295s
4	12	2.6496s
5	9	2.2747s

Tables C.4 and C.5 compare the total linear (inner) iteration numbers and the elapsed CPU times between the modified inexact and exact Newton methods when $u_h^0 = 0.5 * I_h u^*$. Like in Chapter IV, we can see that the modified inexact Newton method is more efficient than the Newton method if the number of unknowns is large.

Table C.4. Comparison between inexact Newton and Newton - Inner iteration numbers in $H_0^1(\Omega)$

h^{-1}	inexact Newton	Newton	ratio (%)
32	50	120	240
64	60	147	245
128	60	184	307
256	65	218	335
512	70	327	284

Table C.5. Comparison between inexact Newton and Newton - Running time in $H_0^1(\Omega)$

h^{-1}	inexact Newton	Newton	ratio (%)
32	0.6259s	0.9799s	157
64	3.0305s	4.8023s	158
128	12.6671s	25.3851s	200
256	59.8509s	133.8117s	224
512	259.5215s	654.9354s	252

CHAPTER VI

CONCLUSIONS

A. Conclusions

In this dissertation, we have applied the modified inexact Newton method to compute approximate solutions to second order nonlinear partial differential equations which have nonlinearities in the highest order derivatives. There were two mesh independent convergence rates for the method: the first one was obtained in the norm of the Sobolev space $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ for $0 < \alpha < 1/2$, the second one in the norm of $H_0^1(\Omega)$.

B. Summary of contributions

In our analysis, it was important to control the norm in $L^\infty(\Omega)$. In Chapter IV, it was bounded by the norm in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ for $0 < \alpha < 1/2$ because the natural inclusion from $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ into $L^\infty(\Omega)$ is continuous in $\mathbb{R}^2$ (see Sobolev embedding theorem). In Chapter V, we used the discrete Sobolev inequalities to bound the norm in $L^\infty(\Omega)$ by the norm in $H_0^1(\Omega)$.

In Chapter IV, we obtained uniform convergence rates of the modified inexact Newton method in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ when Ω is in $\mathbb{R}^2$. First of all, we verified that the nonlinear operator F corresponding to the model problem and its Fréchet derivative F' satisfied the hypotheses, (H.1) and (H.2) in Chapter III. Second, we constructed an iterative scheme to define Newton corrections satisfying the error reduction (A.6) in Chapter III.

In Chapter V, mesh–independent convergence rates of the modified inexact Newton method were given in $H_0^1(\Omega)$ when Ω is in $\mathbb{R}^d$, for $d = 2, 3$. Since convergence results were obtained in the norm of $H_0^1(\Omega)$, we could define Newton corrections using multigrid methods (see, e.g., [11]) or the iterative scheme presented in Chapter IV, setting $\alpha = 0$.

C. Future works

In Chapter IV, we applied the modified inexact Newton method to two dimensional problems in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$. It would be interesting to apply the method to three dimensional problems.

In recent years, several authors (see, e.g., [31, 33, 34, 36]) applied the mixed finite element method to nonlinear problems. In [36], the author computed approximate solutions to mixed problems using Newton's method. It is a future research direction to apply the inexact Newton method to mixed problems.

REFERENCES

[1] I. K. Argyros, *Local convergence of inexact Newton-like iterative methods and applications*, Computers and Mathematics with Applications, 39 (2000), pp. 69–75.

[2] R. E. Bank, *A comparison of two multilevel iterative methods for nonsymmetric and indefinite elliptic finite element equations*, SIAM Journal on Numerical Analysis, 18 (1981), pp. 724–743.

[3] R. E. Bank and T. Dupont, *An optimal order process for solving finite element equations*, Mathematics of Computation, 36 (1981), pp. 35–51.

[4] S. Bellavia and B. Morini, *A globally convergent Newton-GMRES subspace method for systems of nonlinear equations*, SIAM Journal on Scientific Computing, 23 (2001), pp. 940–960.

[5] J. H. Bramble, *On the development of multigrid methods and their analysis*, Proceedings of symposia in applied mathematics, 48 (1994), pp. 5–19.

[6] J. H. Bramble, J. E. Pasciak, and A. H. Schatz, *The construction of preconditioners for elliptic problems by substructuring I*, Mathematics of Computation, 47 (1986), pp. 103–134.

[7] J. H. Bramble, J. E. Pasciak, and P. S. Vassilevski, *Computational scales of Sobolev norms with application to preconditioning*, Mathematics of Computation, 69 (2000), pp. 463–480.

[8] J. H. Bramble, J. E. Pasciak, and J. Xu, *The analysis of multigrid algorithms for nonsymmetric and indefinite elliptic problems*, Mathematics of Computation, 51 (1988), pp. 389–414.

[9] ——, *The analysis of multigrid algorithms with nonnested spaces or noninherited quadratic forms*, Mathematics of Computation, 56 (1991), pp. 1–34.

[10] J. H. Bramble, J. Xu, and J. E. Pasciak, *Parallel multilevel preconditioners*, Mathematics of Computation, 55 (1990), pp. 1–22.

[11] J. H. Bramble and X. Zhang, *The analysis of multigrid methods*, Handbook of Numerical Analysis, 7 (2000), pp. 173–415.

[12] S. C. Brenner, *Discrete Sobolev and Poincaré inequalities for piecewise polynomial functions*, Electronic Transactions on Numerical Analysis, 18 (2004), pp. 42–48.

[13] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, Springer-Verlag, New York, 2002.

[14] P. N. Brown, P. Vassilevski, and C. S. Woodward, *On mesh independent convergence of an inexact Newton-multigrid algorithm*, SIAM Journal on Scientific Computing, 25 (2003), pp. 570–590.

[15] X. Cai and D. E. Keyes, *Nonlinearly preconditioned inexact Newton algorithms*, SIAM Journal on Scientific Computing, 24 (2002), pp. 183–200.

[16] X.-C. Cai and O. B. Widlund, *Multiplicative Schwarz algorithms for some nonsymmetric and indefinite problems*, SIAM Journal on Numerical Analysis, 30 (1993), pp. 936–952.

[17] G. Caloz and J. Rappaz, *Numerical analysis for nonlinear and bifurcation problems*, Handbook of Numerical Analysis, 5 (1998), pp. 487–637.

[18] Z. Cao, *Convergence of multigrid methods for nonsymmetric, indefinite problems*, Applied Mathematics and Computation, 28 (1988), pp. 269–288.

[19] P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland Publishing Company, New York, 1978.

[20] M. Dauge, *Elliptic boundary value problems on corner domains : smoothness and asymptotics of solutions*, Springer-Verlag, New York, 1988.

[21] ——, *Neumann and mixed problems on curvilinear polyhedra*, Integral Equations Operator Theory, 15 (1992), pp. 227–261.

[22] R. S. Dembo, S. C. Eisenstat, and T. Steihaug, *Inexact Newton methods*, SIAM Journal on Numerical Analysis, 19 (1982), pp. 400–408.

[23] P. Deuflhard, *Global inexact Newton methods for very large scale nonlinear problems*, Impact of Computing in Science and Engineering, 3 (1991), pp. 366–393.

[24] ——, *Newton methods for nonlinear problems. Affine invariance and adaptive algorithms*, Springer-Verlag, Berlin, 2004.

[25] S. C. Eisenstat and H. F. Walker, *Globally convergent inexact Newton methods*, SIAM Journal on Optimization, 4 (1994), pp. 393–422.

[26] A. Ern and J.-L. Guermond, *Theory and practice of finite elements*, Springer-Verlag, New York, 2004.

[27] V. Girault, *Finite element methods for Navier-Stokes equations*, Springer-Verlag, New York, 1986.

[28] P. Grisvard, *Elliptic problems in nonsmooth domains*, Pitman, Boston, 1985.

[29] W. Hackbusch, *Multigrid methods and applications*, Springer-Verlag, New York, 1985.

[30] C. Johnson, *Numerical solution of partial differential equations by the finite element method*, Cambridge University Press, Cambridge, 1987.

[31] M.-Y. Kim, F. A. Milner, and E.-J. Park, *Some observations on mixed methods for fully nonlinear parabolic problems in divergence form*, Applied Mathematics Letters, 9 (1996), pp. 75–81.

[32] D. Kincaid and C. S, *Numerical Analysis*, Brooks-Cole Publishing, California, 1996.

[33] F. A. Milner, *Mixed finite element methods for quasilinear second order elliptic problems*, Mathematics of Computation, 44 (1985), pp. 303–320.

[34] F. A. Milner and E.-J. Park, *A mixed finite element method for a strongly nonlinear second–order elliptic problems*, Mathematics of Computation, 64 (1995), pp. 973–988.

[35] B. Morini, *Convergence behavior of inexact Newton methods*, Mathematics of Computation, 66 (1999), pp. 1605–1613.

[36] E.-J. Park, *Mixed finite element methods for nonlinear second–order elliptic problems*, SIAM Journal on Numerical Analysis, 32 (1995), pp. 865–885.

[37] A. Rieder, *On convergence rates of inexact Newton regularizations*, Mathematics of Computation, 66 (1999), pp. 1605–1613.

[38] Y. Saad, *Iterative Methods for sparse linear systems*, PWS Publishing, Boston, 1995.

[39] A. H. Schatz, *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms*, Mathematics of Computation, 28 (1974), pp. 959–962.

[40] V. THOMÉE, *Galerkin finite element methods for parabolic problems*, Springer-Verlag, New York, 1997.

[41] J. XU, *Two-grid discretization techniques for linear and nonlinear PDEs*, SIAM Journal on Numerical Analysis, 33 (1996), pp. 1759–1777.

[42] T. J. YPMA, *Local convergence of inexact Newton methods*, SIAM Journal on Numerical Analysis, 21 (1984), pp. 538–590.

[43] E. ZEIDLER, *Applied functional analysis*, vol. 108 of Applied Mathematical Sciences, Springer-Verlag, New York, 1995.

APPENDIX A

AN ESTIMATE FOR THE RATE OF CONVERGENCE OF THE *W–CYCLE* MULTIGRID METHOD IN $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$

In this appendix, we shall obtain an estimate for the rate of convergence of the *W–cycle* multigrid method applied to the discrete problem corresponding to the Jacobian of the nonlinear PDE introduced in Chapter I. Like in Chapter IV, Ω is a bounded polygonal domain in $\mathbb{R}^2$ and V_h is a finite dimensional subspace of $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$.

We consider the following problem: For a given $f \in V_h$, find $u \in V_h$ such that

$$A(u, v) = (f, v), \quad \text{for all } v \in V_h, \tag{A.1}$$

where $A(u, v) = (F_h'(u^0)u, v)$. Here $u^0 \in V_h$ is an initial iterate of the modified inexact Newton method.

Note that the bilinear form A is nonsymmetric and possibly indefinite. To proceed with our analysis, let $A(u, v) = \hat{A}(u, v) + \hat{D}(u, v)$, where

$$\begin{cases} \hat{A}(u, v) = \displaystyle\int_\Omega k(u^0, x)(\boldsymbol{\nabla} u \cdot \boldsymbol{\nabla} v)\, dx, \\ \hat{D}(u, v) = \displaystyle\int_\Omega \frac{\partial k(u^0, x)}{\partial u^0}(\boldsymbol{\nabla} u^0 \cdot \boldsymbol{\nabla} v) u\, dx + \int_\Omega (\boldsymbol{c} \cdot \boldsymbol{\nabla} u) v\, dx. \end{cases}$$

Then $\hat{A}$ is symmetric and positive definite, and $\hat{D} = A - \hat{A}$.

There are many papers which have estimates for the rate of convergence of multigrid methods applied to symmetric and positive definite problems (see, e.g., [3, 5, 11, 29], and references therein). In contrast, there are a few papers for nonsymmetric and indefinite problems (see, e.g., [2, 8, 11]). For example, in [11], the authors obtained an estimate for the rate of convergence in the norm of $H_0^1(\Omega)$ using a perturbation argument. In this appendix, we shall obtain estimates in the norm of $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ for $0 < \alpha < 1/2$.

First, we shall obtain an estimate for the rate of convergence in the norm of $H_0^{1-\alpha}(\Omega)$. Then, using the discrete inf–sup condition given in Chapter IV, we shall get an estimate in the norm of $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$.

To proceed with our analysis, we assume that $\hat{D}(u,v)$ satisfies the following:

(D.1) For given $0 < \varepsilon < 1$, $|\hat{D}(u,v)| \le Ch^{-\alpha-\varepsilon}\|u\|_{1-\alpha}\|v\|_{1-2\alpha}, \quad$ for all $u, v \in V_h$,

(D.2) $|\hat{D}(u,v)| \le C\|u\|_{1-\alpha}\|v\|_1 \le Ch^{-\alpha}\|u\|_{1-\alpha}\|v\|_{1-\alpha}, \quad$ for all $u, v \in V_h$.

Remark A.1. If $u^0 \in V_h$ is in $H^{1+2\alpha}(\Omega) \cap H_0^1(\Omega)$, that is, $\alpha < 1/4$, then

$$|\hat{D}(u,v)| \le C\|u\|_{1,p}\|u^0\|_{1+2\alpha}\|v\|_{1-2\alpha} \le Ch^{-\alpha-\varepsilon}\|u\|_{1-\alpha}\|v\|_{1-2\alpha},$$

due to Lemma C.1 in Chapter II and inverse inequalities. By (B.2) in Chapter IV, (D.2) holds.

V_h and V_k, for $k = 1, \ldots, J$, are defined as in Section D in Chapter IV. We define operators A_k, $\hat{A}_k$ and $\hat{D}_k : V_k \to V_k$ satisfying

$$(A_k u, v) = A(u,v),\ (\hat{A}_k u, v) = \hat{A}(u,v) \text{ and } (\hat{D}_k u, v) = \hat{D}(u,v),$$

for all $u,\ v \in V_k$. We define smoothing operators $R_k : V_k \to V_k$, for $k = 2, \ldots, J$, such that $R_k = \hat{\lambda}_k^{-1} I$ (the Richardson smoother) where $\hat{\lambda}_k$ is the largest eigenvalue of $\hat{A}_k$. Let Q_k denote the $L^2(\Omega)$ projection onto V_k, for $k = 1, \ldots, J$.

Now we introduce the *W-cycle* multigrid algorithm. Given $f \in V_J$, we shall solve

$$A_J = f.$$

With the initial iterate u^0, we will consider the iterative algorithm

$$u^i = u^{i-1} + B_J(f - A_J u^{i-1}),$$

where $B_j : V_J \to V_J$ is defined by the following multigrid procedure.

Algorithm. Let $p = 2$ and m be a positive integer. Set $B_1 = A_1^{-1}$. Suppose that $B_{k-1} : V_{k-1} \to V_{k-1}$ has been defined. We define $B_k : V_k \to V_k$ as follows. Let $g \in V_k$.

1. (Pre-Smoothing) Set $x^0 = 0$ and define x^l, $l = 1, \ldots, m$ by

$$x^l = x^{l-1} + R_k(g - A_k x^{l-1}).$$

2. (Correction) $y^m = x^m + q^p$, where $q^0 = 0$ and q^i for $i = 1, 2$ is defined by

$$q^i = q^{i-1} + B_{k-1}[Q_{k-1}(g - A_k x^m) - A_{k-1} q^{i-1}].$$

3. (Post-Smoothing) Define y^l for $l = m + 1, \ldots, 2m$ by

$$y^l = y^{l-1} + R_k(g - A_k y^{l-1}).$$

4. $B_k g = y^{2l}$

Let $K_k = I - R_k A_k$, $\hat{K}_k = I - R_k \hat{A}_k$, and the error reduction operator $E_k = I - B_k A_k$ for $k = 1, \ldots, J$. We define an operator $P_k : V_J \to V_k$ satisfying

$$A(P_k u, v) = A(u, v), \quad \text{for all } v \in V_k.$$

Note that if h_k is sufficiently small, then P_k is well defined and satisfies

$$\|(I - P_k)u\|_{1-\alpha} \leq C h_k^{\alpha} \|u\|_1. \tag{A.2}$$

This inequality follows from the result in [39].

For the sake of simplicity, we consider the *W-cycle* multigrid method without post-smoothing. Then, the error reduction operator E_k is given by

$$E_k u = (I - P_{k-1}) K_k^m u + E_{k-1}^2 P_{k-1} K_k^m u, \tag{A.3}$$

where m denotes the number of smoothings.

Since $\hat{A}$ is symmetric and positive definite, we are able to define discrete norms as follows: For $0 \le s < 3/2$,

$$|||u|||_s^2 = (\hat{A}_k^s u, u), \quad \text{for all } u \in V_J. \tag{A.4}$$

Let $\{\phi_i\}_{i=1}^{d_k}$ and $\{\lambda_i\}_{i=1}^{d_k}$ be the orthonormal eigenfunctions and eigenvalues for $\hat{A}_k$, where d_k is the dimension of V_k. If $u = \sum_{i=1}^{d_k} c_i \phi_i \in V_k$, then

$$(\hat{A}_k^s u, u) = \sum_{i=1}^{d_k} \lambda_i^s c_i^2.$$

In fact, $|||u|||_s$ is equivalent to $\|u\|_s$ for all u in V_J. For $0 \le s \le 1$, the equivalence is given in [3]. For $1 < s < 3/2$, since $\hat{A}_k$ is an isomorphism from V_k with the norm $\|\cdot\|_s$ to V_k with the norm $\|\cdot\|_{s-2}$, we have

$$\begin{aligned} \|u\|_s \le C \sup_{v \in V_k} \frac{\hat{A}(u,v)}{\|v\|_{2-s}} &= C \sup_{v \in V_k} \frac{\left(\hat{A}_k^{\frac{s}{2}} u, \hat{A}_k^{\frac{2-s}{2}} v\right)}{\|v\|_{2-s}} \\ \le C \left\|A^{\frac{s}{2}} u\right\|_0 &= C \, ||| \, u \, |||_s \, . \end{aligned}$$

The other inequality follows from

$$\left(\hat{A}_k^s u, u\right) = \hat{A}\left(u, \hat{A}_k^{s-1} u\right) \le C\|u\|_s \left\|A^{s-1} u\right\|_{2-s} \le C\|u\|_s \left(A^s u, u\right)^{1/2}.$$

We then have an estimate for $E_k u$ in the norm $|||\cdot|||_{1-\alpha}$.

Proposition A.2. *Let ε be given in (D.1) and h_0 be sufficiently small. For $0 < h_k < h_0$, if m is large enough, then we can choose a constant $0 < \delta = \delta(m) < 1$*

independent of the mesh size h_k such that

$$\begin{aligned}|||E_k u|||_{1-\alpha} &\leq \{C(1+\delta^2)(m^{-\alpha/2} + h_k^{\alpha-\varepsilon}((1+Ch_k^{\alpha})^m - 1)) + \delta^2(1+Ch_k^{\alpha})^m\} \, |||\, u|||_{1-\alpha} \\ &\leq \delta \, |||\, u|||_{1-\alpha}, \quad \textit{for all } u \in V_k.\end{aligned}$$

Remark A.3. In fact, we can make δ as small as we want increasing m, the number of smoothings.

The proof of the proposition follows the lemma below.

Lemma A.4. *Let ε be given in (D.1). For all $u \in V_k$, we have*

(1) $|||(K_k - \hat{K}_k)u|||_{1-\alpha} \leq Ch_k^{\alpha} \, |||\, u|||_{1-\alpha}$,

(2) $|||K_k u|||_{1-\alpha} \leq (1 + Ch_k^{\alpha}) \, |||\, u|||_{1-\alpha}$,

(3) $|||(K_k - \hat{K}_k)u|||_1 \leq Ch^{\alpha-\varepsilon} \, |||\, u|||_{1-\alpha}$,

(4) $|||\hat{K}_k^m u|||_1 \leq Cm^{-\alpha/2}h_k^{-\alpha} \, |||\, u|||_{1-\alpha}$,

(5) $|||(K_k^m - \hat{K}_k^m)u|||_1 \leq Ch_k^{-\varepsilon}\{(1+Ch_k^{\alpha})^m - 1\} \, |||\, u|||_{1-\alpha}$,

(6) $|||K_k^m u|||_1 \leq \{Cm^{-\alpha/2}h_k^{-\alpha} + Ch_k^{-\varepsilon}((1+Ch_k^{\alpha})^m - 1)\} \, |||\, u|||_{1-\alpha}$,

where C is a generic constant independent of h_k.

Proof. Let $u = \sum_{i=1}^{d_k} c_i \phi_i \in V_k$. By (D.2) and the inverse inequality,

$$\|\hat{D}_k u\|_0 \leq Ch_k^{-1}\|u\|_{1-\alpha}. \tag{A.5}$$

Also we know

$$|||u|||_{1-s} = \sup_{v \in V_k} \frac{\hat{A}(u,v)}{|||v|||_{1+s}},$$

for all $u \in V_k$ and $0 \leq s < 1/2$. By the above inequality,

$$\begin{aligned}
|||(K_k - \hat{K}_k)u|||_{1-\alpha} &= \sup_{v \in V_k} \frac{\hat{A}\left((K_k - \hat{K}_k)u, v\right)}{|||v|||_{1+\alpha}} \\
&= \sup_{v \in V_k} \frac{\left((K_k - \hat{K}_k)u, \hat{A}_k^{\frac{1-\alpha}{2}} v\right)}{\|v\|_0} \\
&= \sup_{v \in V_k} \frac{\frac{1}{\hat{\lambda}_k}\left(\hat{D}_k u, \hat{A}_k^{\frac{1-\alpha}{2}} v\right)}{\|v\|_0} \\
&\leq C h_k^{\alpha} \, ||| \, u \, |||_{1-\alpha} \, .
\end{aligned}$$

The last inequality follows from (A.5), $\hat{\lambda}_k^{-1} \leq C h_k^2$, and

$$\left\| \hat{A}_k^{\frac{1-\alpha}{2}} v \right\|_0 = |||v|||_{1-\alpha} \leq C h_k^{\alpha - 1} \|v\|_0.$$

This completes the proof of (1).

We prove (2) using (1), the triangle inequality, and

$$\begin{aligned}
|||\hat{K}_k u|||_{1-\alpha}^2 &= \left(\hat{A}^{1-\alpha}\left(I - \frac{1}{\hat{\lambda}_k}\hat{A}_k\right) u, \left(I - \frac{1}{\hat{\lambda}_k}\hat{A}_k\right) u\right) \\
&= \sum_{i=1}^{d_k} \lambda_i^{1-\alpha}\left(1 - \frac{\lambda_i}{\hat{\lambda}_k}\right)^2 c_i^2 \leq \sum_{i=1}^{d_k} \lambda_i^{1-\alpha} c_i^2 = |||u \, |||_{1-\alpha}^2 \, .
\end{aligned}$$

The proof of (3) is similar to the proof of (1) using (D.1).

To verify (4), we use the fact

$$\max_{x \in [0,1]} \{x^{\alpha}(1-x)^{2m}\} = \left(\frac{\alpha}{(2m+\alpha)}\right)^{\alpha} \left(\frac{2m}{2m+\alpha}\right)^{2m}.$$

Then,

$$
\begin{aligned}
|||\hat{K}_k^m u|||_1^2 &= \hat{A}(\hat{K}_k^m u, \hat{K}_k^m u) \\
&= \sum_{i=1}^{d_k} \lambda_i \left(1 - \frac{\lambda_i}{\hat{\lambda}_k}\right)^{2m} c_i^2 \\
&= \sum_{i=1}^{d_k} \left(\frac{\lambda_i}{\hat{\lambda}_k}\right)^{\alpha} \left(1 - \frac{\lambda_i}{\hat{\lambda}_k}\right)^{2m} \hat{\lambda}_k^{\alpha} \lambda_i^{1-\alpha} c_i^2 \\
&\leq C h_k^{-2\alpha} m^{-\alpha} \, ||| \, u \, |||_{1-\alpha}^2 \, .
\end{aligned}
$$

Using (2) and (3), (5) is obtained as follows:

$$
\begin{aligned}
|||(K_k^m - \hat{K}_k^m)u|||_1 &\leq |||(K_k - \hat{K}_k)K_k^{m-1}u \, |||_1 + ||| \, \hat{K}_k(K_k^{m-1} - \hat{K}_k^{m-1})u|||_1 \\
&\leq C h_k^{\alpha-\varepsilon} \, ||| \, K_k^{m-1} u \, |||_{1-\alpha} + ||| \, (K_k^{m-1} - \hat{K}_k^{m-1})u|||_1 \\
&\leq C h_k^{\alpha-\varepsilon} \, ||| \, K_k^{m-1} u|||_{1-\alpha} \\
&\qquad + C h_k^{\alpha-\varepsilon} \, ||| \, K_k^{m-2} u \, |||_{1-\alpha} + ||| \, (K_k^{m-2} - \hat{K}_k^{m-2})u|||_1 \\
&\vdots \\
&\leq C h_k^{\alpha-\varepsilon} \sum_{i=1}^{m-1} |||K_k^i u \, |||_{1-\alpha} + ||| \, (K_k - \hat{K}_k)u|||_1 \\
&\leq C h_k^{\alpha-\varepsilon} \sum_{i=0}^{m-1} (1 + C h_k^{\alpha})^i \, ||| \, u|||_{1-\alpha} \\
&\leq C h_k^{-\varepsilon} \{(1 + C h_k^{\alpha})^m - 1\} \, ||| \, u \, |||_{1-\alpha} \, .
\end{aligned}
$$

The triangle inequality completes the proof of (6). □

Proof of Proposition A.2. The proof is given by mathematical induction. Suppose

$$
|||E_{k-1}u|||_{1-\alpha} \leq \delta \, ||| \, u|||_{1-\alpha},
$$

where δ will be determined below. Then, using (A.2), (A.3), and the above lemma gives

$$\begin{aligned}
|||E_k u|||_{1-\alpha} &\le |||(I-P_{k-1})K_k^m u\,|||_{1-\alpha} + |||\,E_{k-1}^2 P_{k-1} K_k^m u|||_{1-\alpha} \\
&\le (1+\delta^2)\,|||\,(I-P_{k-1})K_k^m u\,|||_{1-\alpha} + \delta^2\,|||\,K_k^m u|||_{1-\alpha} \\
&\le C h_k^\alpha (1+\delta^2)\,|||\,K_k^m u\,|||_1 + \delta^2\,|||\,K_k^m u|||_{1-\alpha} \\
&\le \{C(1+\delta^2)(m^{-\alpha/2} + h_k^{\alpha-\varepsilon}((1+Ch_k^\alpha)^m - 1)) \\
&\quad + \delta^2(1+Ch_k^\alpha)^m\}\,|||\,u\,|||_{1-\alpha}\,.
\end{aligned}$$

If m is sufficiently large and h_0 is small enough, then we can choose $0 < \delta < 1$ such that

$$C(1+\delta^2)(m^{-\alpha/2} + h_k^{\alpha-\varepsilon}((1+Ch_k^\alpha)^m - 1)) + \delta^2(1+Ch_k^\alpha)^m \le \delta.$$

This completes the proof of the proposition. □

Remark A.5. In Proposition A.2, for the sake of simplicity, we proved the result when E_k is defined without post-smoothing. The result can be extended to the case involving pre- and post-smoothing.

Let A^T satisfy

$$A^T(u,v) = A(v,u), \quad \text{for all } u, v \in V_J. \tag{A.6}$$

We consider the adjoint problem: For a given f, find $u \in V_J$ such that

$$A^T(u,v) = (f,v), \quad \text{for all } v \in V_J. \tag{A.7}$$

Let E_k^* satisfy

$$A(u, E_k^* v) = A(E_k u, v), \quad \text{for all } u, v \in V_k, \tag{A.8}$$

where E_k is the error reduction operator of the W-*cycle* multigrid method with pre- and post-smoothing for (A.1), that is,

$$E_k u = K_k^m((I-P_{k-1}) + E_{k-1}^2 P_{k-1})K_k^m u.$$

Then, E_k^* is the error reduction operator of the *W-cycle* for the adjoint problem (A.7), that is,

$$E_k^* u = \left(K_k^T\right)^m \left((I - P_{k-1}^*) + (E_{k-1}^*)^2 P_{k-1}^*\right) \left(K_k^T\right)^m u, \text{ for } k = 2, \ldots, J,$$

where P_k^* satisfies $A(P_k u, v) = A(u, P_k^* v)$ and $E_1^* = I - B_1^T A^T = 0$. Furthermore, as in the above proposition, we can choose a sufficiently small constant $0 < \delta_1 = \delta_1(m) < 1$ independent of the mesh size such that

$$|||E_k^* u|||_{1-\alpha} \leq \delta_1 \, ||| \, u|||_{1-\alpha}, \quad \text{for all } u \in V_k. \tag{A.9}$$

Theorem A.6. *Let h_0 be sufficiently small. For $0 < h_k < h_0$, if m is large enough, then there exists a constant $0 < \delta < 1$ independent of the mesh size h_k such that*

$$|||E_k u|||_{1+\alpha} \leq \delta \, ||| \, u|||_{1+\alpha}, \quad \textit{for all } u \in V_k. \tag{A.10}$$

Proof. Using the discrete inf-sup condition in Chapter IV and (A.9), the proof is completed as follows:

$$\begin{aligned}
|||E_k u|||_{1+\alpha} &\leq C \sup_{v \in V_k} \frac{A(E_k u, v)}{|||v|||_{1-\alpha}} \\
&= C \sup_{v \in V_k} \frac{A(u, E_k^* v)}{|||v|||_{1-\alpha}} \\
&\leq C \sup_{v \in V_k} \frac{|||u\,|||_{1+\alpha} \, |||E_k^* v|||_{1-\alpha}}{|||v|||_{1-\alpha}} \\
&\leq C\delta_1 \, ||| \, u|||_{1+\alpha} \\
&\leq \delta \, ||| \, u \, |||_{1+\alpha} \, .
\end{aligned}$$

The last inequality is true because we can choose δ_1 small enough. □

VITA

Taejong Kim was born in Hadong, Korea. He received his B.S. degree in mathematics from Yonsei University in Seoul, Korea in February 1997. From the same university, he earned his M.S. degree in the area of numerical analysis in mathematics in August 2000. He began his Ph.D. degree at Texas A&M University in September 2000 and received it in the area of numerical analysis in May 2006.

Taejong Kim can be contacted by writing to the Department of Mathematics, Texas A&M University, College Station, TX 77843-3368 or to the e–mail address, tjkim@math.tamu.edu.

The typist for this dissertation was Taejong Kim.